Award Number: CDMRP-16-0-PC150752

TITLE: **Development of a Multimarker Urine Test for Prostate Cancer**

PRINCIPAL INVESTIGATOR: **Alvin Y. Liu, PhD**

CONTRACTING ORGANIZATION: **BATTELLE MEMORIAL INSTITUTE**
Richland, WA 99352

REPORT DATE: October 2017

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE **October 2017**	2. REPORT TYPE **Annual**	3. DATES COVERED **30 Sep 2016 - 30 Sep 2017**

4. TITLE AND SUBTITLE
Development of a Multimarker Urine Test for Prostate Cancer

5a. CONTRACT NUMBER

5b. GRANT NUMBER
CDMRP-16-0-PC150752

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Alvin Y. Liu, PhD

E-Mail: aliu@uw.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
   BATTELLE MEMORIAL INSTITUTE
Richland, WA 99352

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Targeted mass spectrometry proteomics was developed (identifying the most suitable surrogate peptides) to measure a set of proteins secreted by cell types of prostate tumors in voided urine. These cancer-associated protein biomarkers were identified by transcriptomic comparison of cancer cells vs. normal luminal cells; cancer-associated stromal cells vs. normal stromal cells. The assays allow quantification of many markers simultaneously in the same biospecimens. Using the levels of PSA for normalization, the multimarker panel achieved an AUC of 0.95 (specificity 1, sensitivity 0.86) in distinguishing cancer from non-cancer. Furthermore, levels of marker MMP9 can be used to distinguish significant cancer from low-risk cancer because MMP9 shows an increased expression in Gleason pattern 4 tumor glands than pattern 3 tumor glands. Biomarker measurements by mass spectrometry and ELISA were in good agreement so that a multiplex ELISA could be developed in the future. Urine collection from patient donors is simple and does not require a digital rectal exam so that donations can be given multiple times with minimal risk since urine is a natural waste product. Antibodies developed for biomarker ELISA can also be used in cancer therapy, especially for a marker (AGR2) that is on the cell surface of cancer cells but inside of normal cells.

15. SUBJECT TERMS
Prostate cancer detection, urine protein biomarkers, targeted mass spectrometry, multimarker panel, detection of significant cancer, tumor-associated surface antigens

16. SECURITY CLASSIFICATION OF: U			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON USAMRMC
a. REPORT Unclassified	**b. ABSTRACT** Unclassified	**c. THIS PAGE** Unclassified	Unclassified	101	**19b. TELEPHONE NUMBER** *(include area code)*

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

Table of Contents

Development of a Multimarker Urine Test for Prostate Cancer

Alvin Y. Liu, PhD

Department of Urology and Institute for Stem Cell and Regenerative Medicine, University of Washington

1. Introduction

Cell types – cancer epithelial and cancer-associated stromal – within prostate tumors have increased expression of genes encoding secreted proteins. These molecules could be detected in voided urine. Targeted multiplex mass spectrometry proteomics was developed for their precise measurement in urine. Protein biomarkers have advantages over, say, RNA biomarkers (e.g., PCA3). Proteins are stable in urine at 4° for days (which allows shipment from collection sites to testing labs), and urine processing does not require the addition of chaotropic agents (to inhibit RNase activities). RNA detection requires amplification, which is labor intensive when larger numbers of samples are tested. Our urine collection is simple and does not place undue stress for the patients. It does not require a DRE (as in the collection for PCA3 testing) so population screening could be carried out. The advantage of multiple marker analysis over single marker analysis in test performance is demonstrated by our results (see recently published paper in *Oncotarget*).

2. Keywords urinary protein biomarkers, voided urine, AGR2 ELISA, multiplex targeted proteomics, prostate cancer detection, multi-marker panel, MMP9, significant cancer detection

3. Accomplishments

Task 1. *Candidate selection for urinary secreted proteins and specific isoforms*
Genes showing elevated tumor expression and encoding secreted/extracellular proteins were identified from cell type-specific transcriptomes. The epithelial-derived markers were AGR2, AGR3, CRISP3, CEACAM5, CEACAM6, CCL3, CCL4, IL24, MMP9; the stromal-derived markers were CXCL14, CD90, IL24, MMP9, POSTN, SFRP4, WISP1. In the *UrinePA* (PeptideAtlas) archive of proteome datasets, the "observed" qualifier (bracketed) is a measure of protein abundance. Of the marker candidates, CRISP3 (65), CEACAM5 (21), CEACAM6 (5), THY1/CD90 (261), MMP9 (115), SFRP4 (17) were listed. Those that were

not listed for healthy donor urine were likely indicative of disease states such as prostate cancer. Other candidates such as KCNG3, KCNH8, LOX, NEO1 showed lower increased expression in cancer.

The existence of variants/isoforms was shown by our PSA analysis. PSA peptides, IVGGWEC$_{cam}$EK and LSEPAELTDAVK, were found to have a correlation coefficient of R^2 = 0.99 across 20 urine samples analyzed. However, data points from samples P08-015C, P07-040C and P06-017C deviated from the correlation plot (Fig. 1). P06-017C urine showed a lower level for LSEPAELTDAVK. Our results indicated an amino acid change in LSEPAELTDAVK. This was supported by a recent discovery of PSA proteoforms, specifically, a nonsynonymous mutation L132I (rs2003783) in LSEPAE(L/I)TDAVK. This novel PSA proteoform was observed in 9 out of 72 clinical serum samples, and P06-017C urine. Abundance recalculation based on this information improved the correlation for this sample. For the other 2 urine samples with lower levels of IVGGWEC$_{cam}$EK, similar type of changes could be the explanation. Thus, any unknown modification in the sequence of surrogate peptides will result in lower protein concentrations. In studies involving human cell lines, most surrogate peptides (453/466) showed a high correlation coefficient (R^2 > 0.8) with no significant data point deviation. The surrogate peptides used in our urine analysis were found to have lower correlation coefficients (R^2 = 0.70) with data point deviations. This observation suggests that
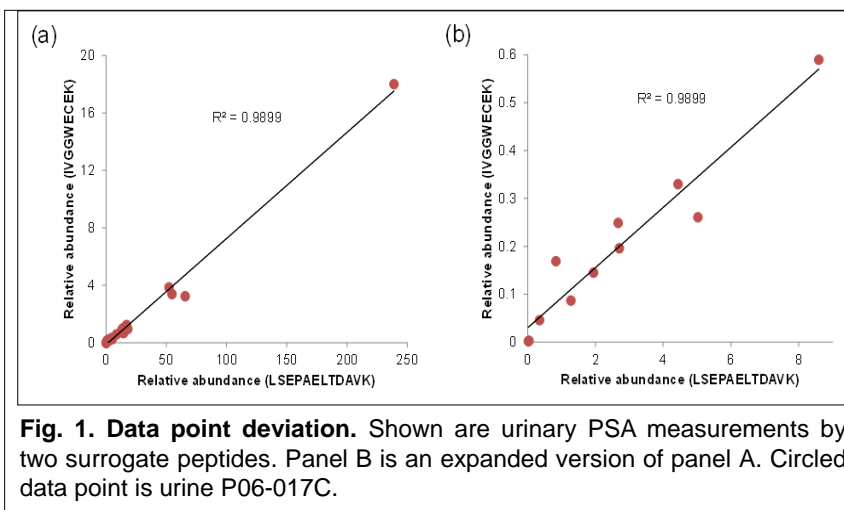

Fig. 1. Data point deviation. Shown are urinary PSA measurements by two surrogate peptides. Panel B is an expanded version of panel A. Circled data point is urine P06-017C.

protein molecules in urine of multiple patient subjects (populations) are less uniform than those in single cell lines (established from individual donors).

Serum [-2]proPSA has been shown to improve performance in combination with PSA for cancer detection. PRISM was able to measure this biomarker in the urine samples. There was not a significant increase in the levels for cancer *vs*. non-cancer urine. Sample P07-031C also yielded the highest value for [-2]proPSA, reflecting the association between [-

2]proPSA and higher Gleason (4+5)/aggressive disease. Other proPSA isoforms ([-5], [-7]), which could be measured simultaneously, appeared to show better performance in distinguishing cancer *vs.* non-cancer.

Task 2. *Development of SRM-based targeted MS assays for candidate proteins and isoforms (including peptide selection and assay optimization)*

Two novel mass spectrometry tools - LG-SRM (long-gradient selected reaction monitoring) and PRISM-SRM (high-pressure high-resolution separation with intelligent selection and multiplexing) - were developed to measure minute quantities of cancer-associated protein biomarkers in body fluids: blood and urine. LG-SRM has ≥10-fold and PRISM-SRM ≥200-fold higher sensitivity than standard LC (liquid chromatography)-SRM. The performance of these tools was demonstrated by our quantification of multiple urinary protein biomarkers in single samples. Selection of the most responsive surrogate peptides of the marker candidates listed above was accomplished through initial screening of 3-5 peptides per protein for the best 5 transitions and their optimal collision energies. For PSA, IVGGWEC$_{cam}$EK and LSEPAELTDAVK were widely used. For the others, a pooled prostate cancer patient urine sample was used to configure the final SRM assays. LG-SRM was used first to measure simultaneously all candidates. CD90, CRISP3, CXCL14, IL24, MMP9, POSTN, and SFRP4 were confidently quantified by at least one peptide. PRISM-SRM was then used to measure AGR2, AGR3, CCL3, CEACAM5, and CEACAM6. The 12 detected markers consisted of 7 moderate-to-low abundance and 5 low abundance proteins (Fig. 2). Undetected CCL4 and WISP1 were excluded from further testing. Performance of
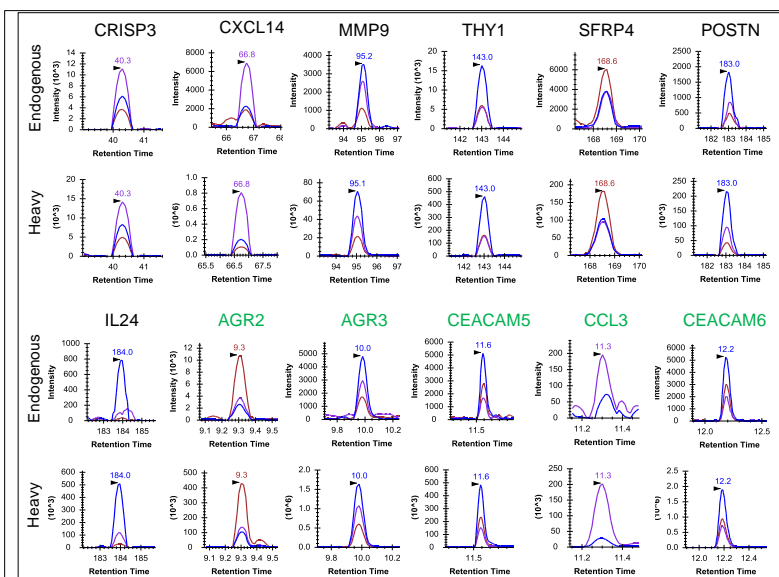

Fig. 2. Signature peptides of biomarkers. Shown are the extracted ion chromatograms of detected proteins in urine sample P07-013C.

the SRM assays was assessed by using a cohort of 14 pre-op and 6 non-cancer urine, and a

cohort of post-op urine (i.e., from men without a prostate). For multiple surrogate peptides of the same protein, a good correlation was achieved when their measurements were in agreement (e.g., R^2 = 0.93 for MMP9 peptides AVIDDAFAR and FQTFEGDLK). Peptides with poor agreement could be due to allelic differences, processing differences, splicing variants, as shown for PSA. For example, the R^2 values for 4 MMP9 peptides ranged from 0.59 between FQTFEGDLK and SLGPALLLLQK to 0.93 between AVIDDAFAR and FQTFEGDLK. A low correlation between VTSLTACLVDQSLR and 2 other peptides was found for CD90, whereas a good correlation, R^2 = 0.72, was found for the other 2 peptides. This suggested that unknown modifications were present in VTSLTACLVDQSLR for several urine samples making this peptide unsuitable for precise CD90 quantitation.

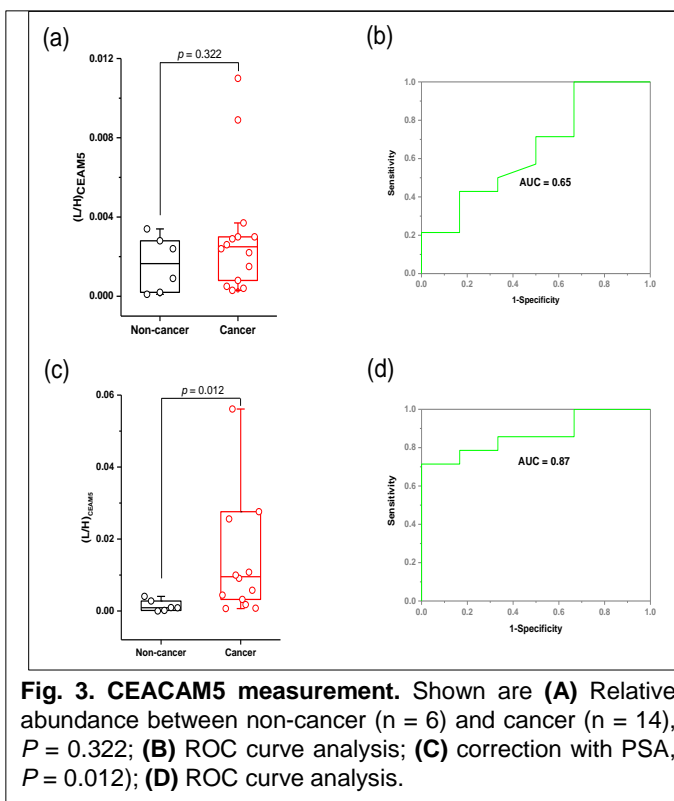**Task 3**. *Urine sample collection and processing*

Over 120 urine specimens, to date, have been collected at UW with non-cancer and different Gleason scores (including a number of Gleason 8-9 cases). The prebiopsy urine were collected under PRoBE design to avoid collection bias. The samples could be either case (biopsy positive) or control (biopsy negative), which was unknown at the time of donation. Pertinent pathology information was requested, which included Gleason score, percentage of Gleason patterns 4 and 5, tumor volume, positive cores. Urine samples from patients under active surveillance for low-risk disease were also obtained. All these specimens were processed by an established protocol (which we have shared with others) involving spin filtration with a molecular weight cutoff of 3K to concentrate urinary proteins from 100 ml to <1 ml in 50 mM NH_4HCO_3. Recently, our proteomics assays were shown to be capable of analyzing proteins from only 5-10 ml of urine. At the same time, RNA from pellet (protein in supernatant) was prepared that would be useful in RNA marker analysis (Quek S *et al.*, Processing of voided urine for prostate cancer RNA biomarker analysis. *Prostate* 2015, **75**:1886-1895; Quek S *et al.*, A multiplex assay to measure RNA transcripts of prostate cancer in urine. *PLoS One* 2012, **9**:e45656). Thus, we have created a biorepository of urinary protein and urinary RNA from the same individual donors. At UW, we are competing with other research groups in the Seattle Prostate Cancer Consortium for urine samples.

Task 4. *Quantitative measurements of secreted protein markers in patient urine samples*

4-1. *Use of PSA levels for marker level normalization* Other organs along the urinary tract could contribute to the protein marker levels (see Task 1). In addition, variations in urinary protein concentration exist among subjects and within subjects. Our experimental data on pre-op *vs.* post-op urine showed that PSA was exclusively secreted from the prostate gland, and urinary PSA levels were proportional to the prostate size. Thus, PSA protein normalization can be used for marker quantification, i.e., for marker levels only contributed by the prostate

Fig. 3. CEACAM5 measurement. Shown are **(A)** Relative abundance between non-cancer (n = 6) and cancer (n = 14), P = 0.322; **(B)** ROC curve analysis; **(C)** correction with PSA, P = 0.012); **(D)** ROC curve analysis.

(tumor). Without PSA normalization, cancer urine showed higher median values than non-cancer urine with most surrogate peptides. Several protein markers showed lower or equal median values in cancer when compared with non-cancer. A Mann-Whitney U test revealed no significant difference between cancer and non-cancer for all the secreted protein markers. With PSA normalization, a significant difference between the cancer and non-cancer was observed with P values ranging from 0.015 to 0.035 (Fig. 3 for CEACAM5). Parenthetically, the cancer-specific urinary PCA3 RNA biomarker score is also generated by normalization with PSA mRNA levels.

4-2. *AUC of 0.95 in differentiating cancer from non-cancer using multiple markers* Receiver operating characteristic (ROC) curve analysis was used to evaluate each surrogate peptide assay in discriminating cancer from non-cancer. The area under the curve (AUC) with 95% confidence intervals showed that peptide assays with P < 0.05 had AUC > 0.80. Furthermore, AUC of well-correlated surrogate peptides were concordant. For example, MMP9 surrogate peptides - AVIDDAFAR, FQTFEGDLK, LGLGADVAQVTGALR, SLGPALLLLQK - yielded AUC of 0.82, 0.86, 0.86, and 0.86, respectively. For surrogate peptides with poorer correlations, lower AUC were obtained.

More importantly, biomarker panel performance by using multivariate analysis of various combinations of the surrogate peptides showed that the best peptide panel was composed of AGR2 (LPQTLSR), AGR3 (LYTYEPR), CEACAM5 (SDLVNEEATGQFR), CD90 (VLYLSAFTSK),
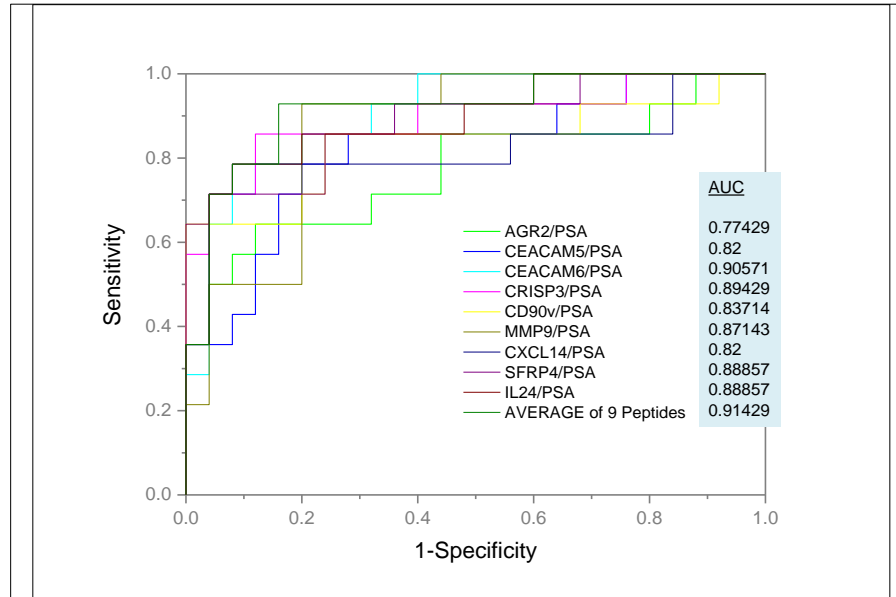

Fig. 4. AUC values of multiple markers in urine analysis. The individual AUC values are listed.

and SFRP4 (GVCISPEAIVTDLPEDVK) with *P* = 0.002 and AUC = 0.95; sensitivity = 0.86 and specificity = 1. Other panels produced AUC values of 0.90 to 0.93 demonstrating the superiority of multi-marker analysis over single-marker analysis. Fig. 4 shows a comparison of the AUC values obtained from single markers and from one marker panel.
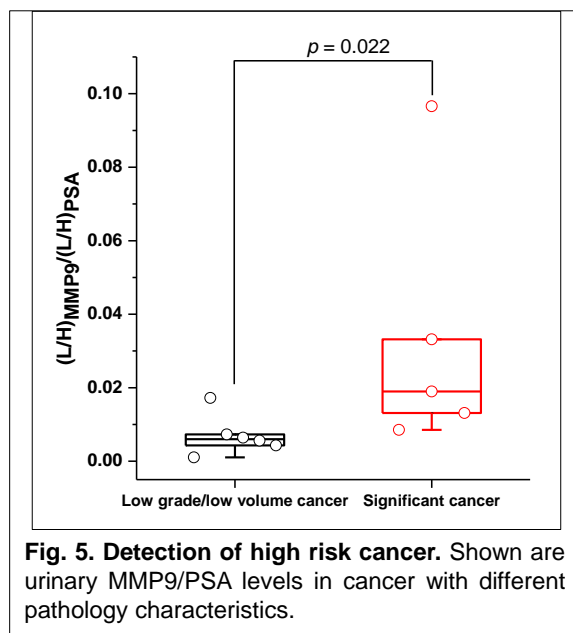
4-3. *Use of MMP9 to identify significant cancer* Our prostate cancer cohort was grouped into low volume/low grade: Gleason score ≤6 and dominant tumor volume ≤ 0.5 cc *vs.* significant (i.e., not meeting the criteria for low volume/low grade disease; Table 1). With most surrogate peptides, the marker/PSA ratios were lower in men with low volume/low grade suggesting that more cancer-associated proteins were produced and secreted into urine from significant cancers. In particular, MMP9 surrogate peptides produced *P* = 0.022 in comparing low volume/low grade cancer and significant cancer with AUC values of >0.90 (Fig. 5). This result was supported by dataset query of cancer cell-type

Fig. 5. Detection of high risk cancer. Shown are urinary MMP9/PSA levels in cancer with different pathology characteristics.

transcriptomes. Array signal intensity value for MMP9 in Gleason pattern 4 cancer cells was 3004.10, ~12-fold higher than that of 238.41 in Gleason pattern 3 cancer cells. For comparison, urine PSA (P = 0.93) and serum PSA (P = 0.32) in this cohort showed no power in differentiating the two cancer types.

In summary, our proteomics results validated our cell type-specific transcriptomics results from which these biomarker candidates were identified. There was a good correlation between protein expression and mRNA expression in contrast to reports in the literature stating no significant correlation between the two.

Task 5. *Quantitative measurement of secreted proteins in patient urine using ELISA*

ranking	case	Gleason	tumor volume	stage
1	P06-003	3+3 bx	na	na
2	P06-017	3+3 bx	na	na
3	P07-018	3+3 bx	na	na
4	P07-040	3+3	0.1	T2a
5	P07-029	3+3	0.5	T2c
6	P08-006	3+3	0.5	T2c
7	P08-015	3+4	2.5	T3b
8	P08-028	3+4	4	T2c
9	P08-032	4+3	1	T2c
10	P07-047	4+3	1.5	T2c
11	P07-031	4+5 bx	na	TxN1

Table 1. Cancer severity ranking. bx = biopsy, na = not available, tumor volume in cc.

5-1. *Correlation between AUC values obtained by targeted SRM and ELISA* The urine samples were also analyzed by ELISA for AGR2. The AUC value obtained was 0.74, which matched well the value of 0.77 obtained from PRISM-SRM. Hence, both biomarker measurement methods produced concordant marker quantities (R^2 = 0.91). The AUC for AGR2 was affected by "non-cancer" AGR2 due to the presence of AGR2-expressing organs along the
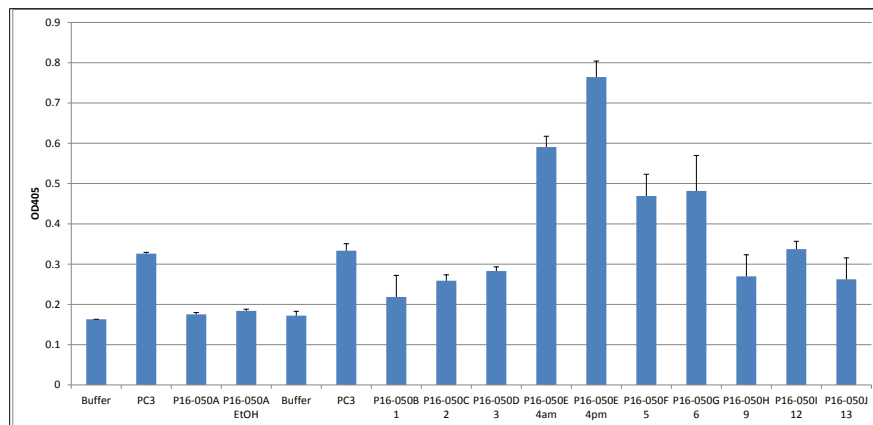

Fig. 6. AGR2 in non-cancer urine. Urine collected in 14 days were analyzed. An increase was seen in P16-050E (both morning and afternoon donations) followed by decline on subsequent days. Alcohol lysis of cells in urine did not increase detectable AGR2 (bars 3 *vs.* 4). Media of prostate cancer cell line PC3 is the positive control.

urinary tract, e.g., bladder urothelium, kidney tubules (moderately stained for AGR2). A fraction of non-cancer urine was found to have above background amount of urinary AGR2, although urine proteome database query yielded

no AGR2 identifier. AGR2 levels in a middle age non-cancer male (with normal sPSA) were

measured over a two-week period. There was an increase on day 4 (for both morning and afternoon urine) followed by a decrease to baseline level on subsequent days (Fig. 6). Treatment of urine sample by the addition of alcohol (to lyse shed bladder cells) did not raise the level of detectable urinary AGR2. Thus, day-to-day physiological differences could account for AGR2 in the urine of non-cancer. Two female urine were also showed measurable amounts of AGR2. Unlike non-cancer, AGR2 in cancer urine would remain above background, and the urine would contain other marker proteins. Testing this hypothesis would require a suitable active surveillance patient donor to volunteer donations every day over a similar period of two weeks.

5-2. *Discovery of cell surface expression of AGR2 specific to cancer* Our collaborator at the University of London have shown that our developed antibodies to AGR2, P1G4 and P3A5, could target (Agr2+ mouse pancreatic) tumor cells in mice (all anti-human AGR2 antibodies tested recognize the murine Agr2 as the two proteins are evolutionarily conserved). Significantly, no normal tissue showed (radiolabeled) antibody uptake. AGR2 in normal cells are found in the endoplasmic reticulum where it functions as a protein disulfide isomerase. The cell interior localization makes normal cells immune to AGR2 antibody. This cancer-specific subcellular localization predicts a strong likelihood that anti-AGR2 will be highly effective against AGR2-positive tumor cells (which include pancreatic, prostate, breast, lung, oral) without significant collateral damage to healthy cells. Another supportive argument is in the generation of AGR2 monoclonals. As the generated antibodies could recognize mouse Agr2, one would expect severe deleterious effect on the animals' health with abundant circulating anti-AGR2/Agr2 post-immunization. At the time of sacrifice, the animals appeared healthy, and many of the organs examined showed no evidence of cytotoxicity. That normal cells do not secrete AGR2 is supported by the finding of background level of AGR2 in serum using PRISM-SRM. No Agr2 secretion by normal tissue in mice was indicated by the correlation between serum Agr2 level and implanted tumor size. The bladder urothelium did not secrete appreciable amounts of AGR2 into the urine (Ho ME *et al.*, Bladder cancer cells secrete while normal bladder cells express but do not secrete AGR2. *Oncotarget* 2016, **7**:15747).

Currently, we are measuring the marker levels in the large cohort of samples to validate the AUC obtained. Additional samples could be provided by the Center for Prostate Disease Research and the University of Texas at San Antonio.

Nothing to report on opportunities for training and professional development; on results disseminated to communities of interest.

4. Impact

A urine test with high performance characteristics (AUC = 0.95, specificity = 1) can likely be used to replace the physical biopsy to diagnose prostate cancer, which is invasive, expensive, and potentially morbid with up to a 4% risk of sepsis. In contrast, the urine test is non-invasive and relatively cheap, and poses minimal risk to the patients. The serum PSA test (AUC = ~0.6 to 0.7) is imperfect in that a majority of patients with abnormal values turn out to have no cancer detected by biopsy (biopsy negative urine samples). The high biopsy numbers impose a severe strain on our health care. More importantly, urinary MMP9 measurement can be used to distinguish significant cancer, which no other tests can consistently. Since urine donation is risk free, urine testing can be used to monitor patients on active surveillance more frequently for progression than periodic biopsies.

Reagents developed for biomarker analysis such as highly specific monoclonal antibodies can be used for cancer immunotherapy, especially when cancer cells differ from normal cells in the localization of the antigen: cell surface *vs*. cell interior.

The proteomics tools developed can be employed for other cancer types. For example, AGR2 is overexpressed and secreted by many solid tumor types including the major ones of lung (Alavi M *et al.*, High expression of AGR2 in lung cancer is predictive of poor survival. *BMC Cancer* 2015, **15**:655) breast, colorectal, pancreatic. As described, our anti-AGR2 antibodies can be used for both prostate and pancreatic cancers. AGR2 is found in most prostate distal metastases (Ho ME *et al.*, Prostate cancer cell phenotypes based on AGR2 and CD10 expression. *Mod Pathol* 2013, **26**:849-859).

Nothing to report on the impact on technology transfer; on society beyond science and technology.

5. Changes/Problems Nothing to report

6. Products

Journal publications: Shi T, Quek S, Gao Y, Nicora CD, Nie S, Fillmore TL, Liu T, Rodland KD, Smith RD, Leach RJ, Thompson IM, Vitello EA, Ellis WJ, Liu AY, Qian W. Multiplexed targeted mass spectrometry assays for prostate cancer-associated urinary proteins. *Oncotarget* 2017, Epub. Supported in part by PCRP W81XWH-16-1-0614.

Books or other non-periodical, one-time publications: none to report.

Other publications, conference papers, and presentations: none to report.

Websites or other Internet sites: none to report.

Patent applications: non-provisional application (Klarquist Ref. No. 23-97721-02) based on a provisional application filed in 2016 (No. 62/432,946; Battelle Ref. No. 309999-E PROV) on Panel of Prostate Cancer-associated Secreted Proteins for Diagnosis and Prognosis for the work carried out by UW and Pacific Northwest National Laboratory.

Other Products: biorepositories of urinary protein and urinary RNA for biomarker analysis.

7. Participants & Collaborating Organizations

Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA99352 is our partnering institute in this research:

Wei-Jun Qian - PI, Tujin Shi - research scientist, Carrie D Nicora - research scientist

Active support of PI – in the last reporting period, a previously active grant (NCI CA111244) on Biomarker Development has closed and the PI's research assistant (Elizabeth Vitello) has gone back to graduate school.

There is nothing to report on other organizations. Two other institutions are sending us their urine samples for testing as identified above.

8. Special Reporting Requirements

Our collaborating partner has been sending in quarterly reports.

Collaborative Awards: nothing to report.

Quad Charts: nothing to report.

9. Appendices

1. Shi T, Quek S, Gao Y, Nicora CD, Nie S, Fillmore TL, Liu T, Rodland KD, Smith RD, Leach RJ, Thompson IM, Vitello EA, Ellis WJ, Liu AY, Qian W. Multiplexed targeted mass spectrometry assays for prostate cancer-associated urinary proteins. *Oncotarget* 2017, Epub.

2. Non-provisional US patent application: *Panel of Prostate Cancer-associated Secreted Proteins for Diagnosis and Prognosis*.

Multiplexed targeted mass spectrometry assays for prostate cancer-associated urinary proteins

Tujin Shi[1], Sue-Ing Quek[2,3,*], Yuqian Gao[1], Carrie D. Nicora[1], Song Nie[1], Thomas L. Fillmore[4], Tao Liu[1], Karin D. Rodland[1], Richard D. Smith[1], Robin J. Leach[5], Ian M. Thompson[5], Elizabeth A. Vitello[2,3], William J. Ellis[2], Alvin Y. Liu[2,3] and Wei-Jun Qian[1]

[1]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA

[2]Department of Urology, University of Washington, Seattle, WA

[3]Institute for Stem Cell and Regenerative Medicine, University of Washington, Seattle, WA

[4]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA

[5]Department of Urology and the Cancer Therapy and Research Center, University of Texas Health Science Center at San Antonio, San Antonio, TX

[*]Present address: Singapore Polytechnic, Center for Biomedical and Life Sciences T11A-412 (level 4), Singapore

Correspondence to: Wei-Jun Qian, email: weijun.qian@pnnl.gov
Alvin Y. Liu, email: aliu@uw.edu

Keywords: secreted protein biomarkers, prostate cancer detection, prostate cancer, targeted mass spectrometry, selected reaction monitoring

Received: May 05, 2017   Accepted: August 04, 2017   Published: October 09, 2017

ABSTRACT

Biomarkers for effective early diagnosis and prognosis of prostate cancer are still lacking. Multiplexed assays for cancer-associated proteins could be useful for identifying biomarkers for cancer detection and stratification. Herein, we report the development of sensitive targeted mass spectrometry assays for simultaneous quantification of 10 prostate cancer-associated proteins in urine. The diagnostic utility of these markers was evaluated with an initial cohort of 20 clinical urine samples. Individual marker concentration was normalized against the measured urinary prostate-specific antigen level as a reference of prostate-specific secretion. The areas under the receiver-operating characteristic curves for the 10 proteins ranged from 0.75 for CXL14 to 0.87 for CEAM5. Furthermore, MMP9 level was found to be significantly higher in patients with high Gleason scores, suggesting a potential of MMP9 as a marker for risk level assessment. Taken together, our work illustrated the feasibility of accurate multiplexed measurements of low-abundance cancer-associated proteins in urine and provided a viable path forward for preclinical verification of candidate biomarkers for prostate cancer.

INTRODUCTION

Prostate cancer is the most common solid tumor in men and the second leading cause of male cancer-related deaths in the US. Over-diagnosis and over-treatment of prostate cancer have become major concerns for disease management ever since the introduction of serum prostate-specific antigen (sPSA) screening [1, 2]. There is still a significant need to develop informative biomarkers for effective non-invasive detection of high risk prostate cancer, the ones that need to be treated, from the many low risk non-life threatening cancer cases.

Human urine is an ideal clinical specimen for testing prostate cancer biomarkers since prostatic secretion passes

into the urine. Currently, one prostate cancer urine test measures a cancer-specific non-coding transcript PCA3 released from prostate cancer cells [3]. In a cohort of >500 patients with serum PSA between 3 and 15 ng/mL, the area under the receiver-operating characteristic curve (AUC) was 0.66 with a sensitivity of 65% and a specificity of 66% [4]. As a prognostic marker, PCA3 showed no significant link to Gleason score, tumor volume, and cancer stage in a cohort of 70 cases [5], though a link to tumor volume and surgical margin was reported in another study [6]. PCA3 is a low abundance transcript, and an "attentive" digital rectal exam (DRE) by an experienced urologist is required to enhance the PCA3 signal [7]. Since most current clinical tests are based on protein analytes, there is an interest in identifying better protein biomarkers for prostate cancer. Moreover, proteins are more stable than RNA, which requires the addition of a preservative to the urine sample and immediate processing.

We have previously identified a set of prostate cancer-associated secreted protein markers by cell-type transcriptomics [8, 9] for quantification in urine. Assay developments for measuring single secreted protein markers in voided urine have been reported [8-12]. For example, AGR2 (anterior gradient 2) is produced in relatively high abundance by cancer epithelial cells [9]. Compared with benign tissue, AGR2 is highly expressed in tumors at the mRNA and protein levels [10, 13]. A sandwich ELISA and a highly sensitive targeted mass spectrometric approach termed PRISM (high-pressure, high-resolution separation with intelligent selection and multiplexing) coupled with selected reaction monitoring (SRM) were used to measure AGR2 in human urine at pg/mL levels [11]. We demonstrated that the amounts of urinary AGR2 measured by both ELISA and PRISM-SRM in the same samples were concordant with $R^2 = 0.91$. Our initial cohort study indicated that urinary AGR2 was able to differentiate prostate cancer from non-cancer urine with an AUC = 0.75 [11].

Herein, we report multiplexed measurements of 12 cancer-associated proteins in urine by targeted mass spectrometry (MS) and the potential utility of these markers for prostate cancer detection. SRM-based targeted MS has proven to be a reliable technology for accurate quantification of target proteins due to its high reproducibility, multiplexing, and specificity whereas antibodies can sometimes show unexpected cross reactivity [14, 15]. A major limitation of typical liquid chromatography (LC)-SRM analysis is the insufficient sensitivity to detect low-abundance proteins in body fluids (e.g., <1 ng/mL in blood plasma/serum), encountered as in early detection [14]. We recently introduced two highly sensitive complementary targeted proteomics approaches: long gradient (LG)-SRM [16] and PRISM-SRM [17, 18] for reliable detection and quantification of low-abundance proteins in body fluids and human tissues. LG-SRM and PRISM-SRM were demonstrated to provide ≥10-fold and

≥200-fold higher sensitivity, respectively, when compared to standard LC-SRM. To enable multiplexed quantification of prostate cancer associated protein markers in urine, we have developed sensitive SRM assays for direct detection of these markers in voided urine without entailing DRE. The multiplexed SRM assays provide a means for verifying the performance of individual markers or multi-marker panel for prostate cancer detection. Once promising markers are identified and verified in initial cohort studies, antibody-based ELISA assays can be developed for high-throughput clinical applications.

RESULTS

Tumor-associated secreted proteins in human urine

Through comparison of cell type-specific transcriptomes, genes showing elevated tumor expression and encoding secreted/extracellular proteins were identified from both the epithelial and stromal compartments. Furthermore, gene expression analysis indicated that many showed differential expression among tumors of different Gleason scores. The epithelial derived marker candidates included AGR2, AGR3, CRISP3, CEAM5, CEAM6, CCL3, CCL4, IL24, MMP9; the stromal derived candidates included CXL14, CD90, IL24, MMP9, POSTN, SFRP4, and WISP1. In the *UrinePA* (peptide atlas, http://www.peptideatlas.org) archive of proteome datasets, the "observed" (in brackets) qualifier was used to indicate protein abundance. Of the marker candidates, CRISP3 (65), CEAM5 (21), CEAM6 (5), CD90/*THY1* (261), MMP9 (115), SFRP4 (17) were listed (Supplementary Table 1). Those that were not detected in healthy donors could be either below the limit of detection or likely specific for disease (e.g., prostate cancer).

Multiplexed SRM assays for prostate cancer protein markers

To develop targeted SRM assays for individual proteins, selection of the most suitable surrogate peptides for each protein was critical for precise quantification of target proteins in patient specimens. The initial selected surrogate peptides for each protein marker are listed in Supplementary Table 2. The peptide selection follows several main criteria: a) sequences being unique to their corresponding proteins; b) peptides having high MS response and minimal matrix interference in LC-SRM signals; c) generally no known modifications or mutations within the selected peptide sequences.

For PSA, IVGGWECEK and LSEPAELTDAVK were demonstrated to be the most effective [17, 19]. For the others, a pooled prostate cancer patient urine sample was used to configure the final SRM assays with evaluation of matrix interference, endogenous peptide

detectability and peptide SRM response. LG-SRM was used first to measure all candidates simultaneously due to its moderate sensitivity ($\geq$ 10-fold higher than LC-SRM) and higher multiplexing capability (~3 times higher than LC-SRM) [16]. PSA, CD90, CRISP3, CXL14, IL24, MMP9, POSTN, and SFRP4 were confidently detected and quantified by at least one surrogate peptide (Figure 1 and Table 1). More sensitive PRISM-SRM ($\geq$20-fold higher in sensitivity than LG-SRM [17]) was then used to measure the remainder. AGR2, AGR3, CCL3, CEAM5, and CEAM6 were reliably detected and quantified except CCL4 and WISP1 (Figure 1). The reproducibility of LG-SRM and PRISM-SRM based assays for measurements in biofluids such as urine and serum was well validated in our previous reports, which typically had a coefficient of variance (CV) <10% [16, 17, 20]. With a combined LG-SRM and PRISM-SRM, SRM assays were established for each of the detectable peptides: three best transitions without matrix interference and the best transition for quantitation (Table 1). We note that two peptides, LYTYEPR for AGR3 and MVIITTK for CXL14, may not serve as good surrogates for protein quantification because of the reported phosphorylation sites as well as the potential oxidation on the methionine residue for MVIITTK.

From the assay results, the 12 detected markers were grouped into 7 moderate-to-low abundance proteins for LG-SRM and 5 low abundance proteins for PRISM-SRM. CCL4 and WISP1 were excluded from further testing. The SRM assays were then applied for marker quantification in a cohort of 14 cancer (pre-op) and 6 non-cancer (healthy control) urine collected at the University of Washington (UW), and a cohort of post-op urine collected at the University of Texas Health Science Center at San Antonio (UTHSCSA) for urinary PSA contribution by the prostate. Among the 12 proteins, 10 proteins can be reliably detected and quantified across the 20 urine subjects with at least one surrogate peptide, except CCL3 and POSTN.

Concordance between multiple surrogate peptides from the same protein

Since we selected multiple surrogate peptides for quantification of a specific protein in urine, we

Figure 1: Extracted ion chromatograms (XICs) of detected proteins in a single urine sample, P07-031C. Seven proteins (CD90, CRISP3, CXL14, IL24, MMP9, POSTN, SFRP4) were detected by LG-SRM, and the other five (AGR2, AGR3, CCL3, CEAM5, CEAM6) in extremely low abundance were detected by PRISM-SRM. Three transitions (blue, chestnut, and purple curves) for one surrogate peptide of each protein were monitored. The surrogate peptides being monitored are: CD90 (*THY1*): VLYLSAFTSK, CRISP3: WANQC_cam_NYR, CXL14: MVIITTK, IL24: LWEAFWAVK, MMP9: AVIDDAFAR, POSTN: AAAITSDILEALGR, SFRP4: GVC_cam_ISPEAIVTDLPEDVK, AGR2: LPQTLSR, AGR3: LYTYEPR, CCL3: QVC_cam_ADPSEEWVQK, CEAM5: SDLVNEEATGQFR, CEAM6: SDPVTLNVLYGPDGPTISPSK.

Table 1: Prostate cancer-associated secreted proteins and their surrogate peptides

Protein	Accession number	Best surrogated peptide[a]	SRM transitions			
			Q1		Q3	
AGR2	O95994	LPQTLSR	407.7	604.3	476.3	351.2
AGR3	Q8TD06	LYTYEPR	471.2	665.3	272.2	277.2
CCL3	P10147	QVCADPSEEWVQK[b]	788.4	1188.6	1117.5	1002.5
CEAM5	P06731	INGIPQQHTQVLFIAK	603.0	847.5	761.9	705.4
		SDLVNEEATGQFR	733.3	1051.5	937.4	679.4
		CETQNPVSAR[b]	581.3	872.5	643.4	529.3
CEAM6	P40199	EVLLLAHNLPQNR	506.3	741.4	514.3	531.8
		SDPVTLNVLYGPDGPTISPSK	1079.1	1055.5	998.5	331.2
CRISP3	P54108	WANQCNYR[b]	556.2	925.4	854.4	612.3
		YEDLYSNCK[b]	596.3	899.4	784.4	671.3
CXL14	O95715	MVIITTK	403.2	674.4	575.4	462.3
		WYNAWNEK	555.8	761.4	647.3	576.3
IL24	Q13007	LWEAFWAVK	575.3	850.4	721.4	650.4
MMP9	P14780	AVIDDAFAR	489.3	807.4	694.3	579.4
		FQTFEGDLK	542.8	809.4	708.4	561.3
		LGLGADVAQVTGALR	720.9	914.5	815.5	744.4
		SLGPALLLLQK	576.9	952.6	727.5	614.4
POSTN	Q15063	AAAITSDILEALGR	700.9	1074.6	973.5	771.5
SFRP4	Q6FHJ7	GVCISPEAIVTDLPEDVK[b]	971.5	1425.7	916.5	587.3
CD90	P04216	VLYLSAFTSK	564.8	916.5	753.4	640.3
		VTSLTACLVDQSLR[b]	521.6	830.5	717.4	618.3
		HVLFGTVGVPEHTYR	571.3	958.5	802.4	576.3

[a]These surrogate peptides were confidently detected in the pooled urine sample.
[b]Cysteine was synthesized as carbamidomethyl cysteine.

evaluated the agreement between these peptides from the same protein. Conceptually, when no posttranslational modifications or undocumented amino acid changes exist in the surrogate peptides, their measured concentrations should have a high degree of correlation across all samples because the surrogate peptide level was stoichiometric to that of their cognate protein [21]. With any peptide sequence modifications, the level of the unmodified surrogate peptides would be lower, affecting accurate measurement of their corresponding proteins. Given the possibility of unknown sequence modifications, each surrogate peptide could potentially represent a distinctive signature with diagnostic value [22]. To evaluate the quantification accuracy, correlation analysis of the L/H ratios between the surrogate peptides from the same protein was carried out. For example, MMP9 was represented by four quantifiable surrogate peptides,

and the Pearson correlation coefficients ranged from 0.59 for FQTFEGDLK and SLGPALLLLQK to 0.93 for AVIDDAFAR and FQTFEGDLK, which suggested that multiple MMP9 isoforms could exist in these clinical urine samples (Figure 2). For CD90, low correlation coefficients between VTSLTACLVDQSLR and two other peptides were obtained, whereas a good correlation, $R^2 = 0.72$, was found for the other two peptides (Supplementary Figure 1). This suggested the presence of unknown modifications in VTSLTACLVDQSLR in several urine samples, making this peptide unsuitable for accurate measurement of CD90.

The origin of urinary PSA

To assess whether urinary PSA is exclusively originating from the prostate, LC-SRM was used to measure its concentrations in 7 urine samples from men

Table 2: Performance of surrogate peptide markers derived from 10 prostate cancer-associated secreted proteins in 20 urine samples (14 cancer and 6 non-cancer samples)

Protein	Peptide	(L/H)$_{peptide\ marker}$		(L/H)$_{peptide\ marker}$/(L/H)$_{PSA}$			
		P value[a]	AUC	*P* value[a]	AUC	Sensitivity[b]	Specificity[b]
AGR2	LPQTLSR	0.773	0.45	0.063	0.77	0.93	0.67
AGR3	LYTYEPR	0.283	0.66	0.019	0.85	0.79	1
CEAM5	SDLVNEEATGQFR	0.322	0.65	0.012	0.87	0.71	1
CEAM6	EVLLLAHNLPQNR	0.246	0.67	0.029	0.82	0.79	0.83
CRISP3	WANQCNYR[c]	0.386	0.63	0.035	0.86	0.86	0.83
CRISP3	YEDLYSNCK[c]	0.433	0.38	0.035	0.81	0.64	1
CD90	VLYLSAFTSK	0.174	0.70	0.015	0.86	0.86	0.83
CD90	VTSLTACLVDQSLR[c]	0.967	0.45	0.063	0.77	0.64	1
CD90	HVLFGTVGVPEHTYR	0.650	0.57	0.012	0.87	0.86	0.83
CXL14	MVIITTK	0.836	0.46	0.091	0.75	0.79	0.83
IL24	LWEAFWAVK	0.479	0.61	0.015	0.86	0.71	1
MMP9	AVIDDAFAR	1	0.50	0.029	0.82	0.93	0.67
MMP9	FQTFEGDLK	0.710	0.56	0.015	0.86	0.93	0.67
MMP9	LGLGADVAQVTGALR	0.869	0.47	0.015	0.86	0.86	0.83
MMP9	SLGPALLLQK	1	0.49	0.015	0.86	0.85	0.83
SFRP4	GVCISPEAIVTDLPEDVK[c]	0.592	0.42	0.023	0.83	0.64	1

[a]*P* values were obtained from the Mann-Whitney U test.
[b]These are the sensitivity and specificity at the optimal cutoff point (i.e., the best sum of sensitivity and specificity).
[c]Cysteine was synthesized as carbamidomethyl cysteine.

after radical prostatectomy (i.e., the entire prostate being removed) and the cohort of 20 urine samples before radical prostatectomy (Supplementary Table 3). The measured PSA levels ranged from 0.02 ng/100 μg to 2.95 ng/100 μg of total protein with an average value of 0.98 ng/100 μg (median 0.41 ng/100 μg, Supplementary Table 4). When compared with the PSA levels in the others with an average value of 110.89 ng/100 μg of total urinary protein (median 28.68 ng/100 μg), the PSA percentage in the post-op urine was ~1% (median ~1.5%, Supplementary Table 4). Thus, our data showed that urinary PSA was secreted exclusively from the prostate, and the contribution from other sources in the urinary system was negligible.

Initial assessment of marker utility in a pilot cohort

In SRM measurements, the L/H peak area ratios were proportional to the concentrations of their cognate protein, which were expressed as ng/100μg of total urinary protein because of the same peptide concentration with the same amount of spiked-in heavy internal standards (see Supplementary Methods). Thus, the L/H ratio could be regarded as the adjusted concentration of the target protein in urine (against the total amount of urinary proteins [11], Supplementary Table 5). This adjustment accounted for a substantial degree of variations in urinary protein concentration among donors, and donations from the same donor. For most surrogate peptides measured, the cancer urine showed higher median L/H values than non-cancer urine; while for several others (CRISP3, CXL14, IL24 and SFRP4), a lower or equal median L/H value in cancer *vs*. non-cancer was found. A Mann-Whitney U test of the surrogate peptide L/H ratios revealed no significant difference between cancer and non-cancer for all the markers (Table 2).

Since prostate cancer associated proteins are mostly secreted from the prostate tissue, we considered a "normalization" strategy against a baseline level of prostate specific secretion. For this purpose, we adapted the strategy to normalize all marker concentrations against urinary PSA concentration. We chose urinary PSA level as a reference value of prostate specific secretion because our data showed that urinary PSA was exclusively secreted from the prostate gland. Similar normalization strategy was applied in the urine PCA3 assay where the marker score was generated by normalization of the PCA3 transcript levels to those of PSA transcript [23].

The protein marker/PSA concentration ratios were obtained by dividing the L/H peak area ratio of surrogate marker peptides by that of PSA peptide IVGGWEC$_{cam}$EK (Supplementary Table 6). After PSA normalization, a significant difference between the cancer and non-cancer urine was observed for the marker peptides [except for LPQTLSR of AGR2, VTSLTACLVDQSLR of CD90 and MVIITTK of CXL14] with $P = 0.015$-0.035 (Table 2 and Figure 3). ROC analysis with 95% confidence intervals showed that the peptides with $P < 0.05$ produced AUC values >0.80, while for the three peptides with $P > 0.05$ the AUC values produced were <0.80 (Table 2). These
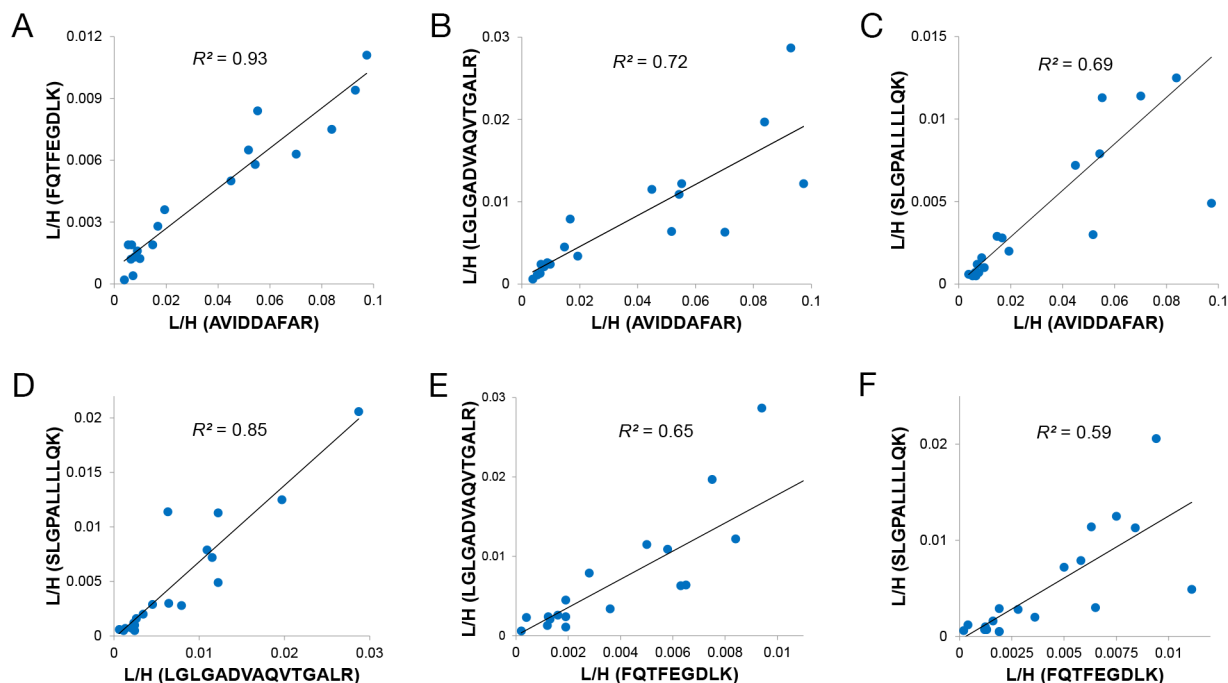

Figure 2: Correlation plot between any two MMP9 surrogate peptides in 20 urine samples. (A) Relative abundance correlation between FQTFEGDLK (*y*-axis) and AVIDDAFAR (*x*-axis); **(B)** Relative abundance correlation between LGLGADVAQVTGALR (*y*-axis) and AVIDDAFAR (*x*-axis); **(C)** Relative abundance correlation between SLGPALLLLQK (*y*-axis) and AVIDDAFAR (*x*-axis); **(D)** Relative abundance correlation between SLGPALLLLQK (*y*-axis) and LGLGADVAQVTGALR (*x*-axis); **(E)** Relative abundance correlation between LGLGADVAQVTGALR (*y*-axis) and FQTFEGDLK (*x*-axis); **(F)** Relative abundance correlation between SLGPALLLLQK (*y*-axis) and FQTFEGDLK (*x*-axis). L/H = the ratio of SRM signal from endogenous peptide over heavy-labeled internal standard. R^2 values range from 0.59 to 0.93.
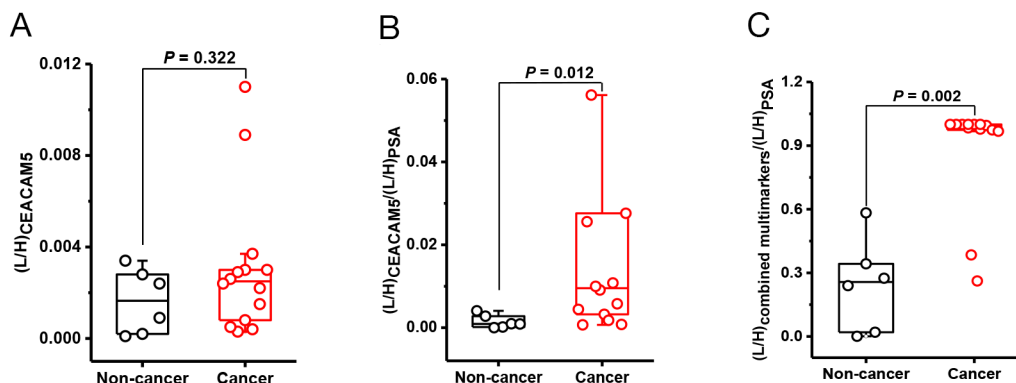

Figure 3: Urine protein biomarkers for prostate cancer. (A) CEAM5 relative abundance between non-cancer (n = 6) and cancer urine (n = 14), $P = 0.322$; **(B)** CEAM5/PSA concentration ratios between non-cancer and cancer, $P = 0.012$; **(C)** Significant differentiation between non-cancer and cancer, $P = 0.0034$, with the best peptide combination. The relative abundance of CEAM5 and PSA was derived from their surrogate peptides, SDLVNEEATGQFR and IVGGWEC$_{cam}$EK, respectively. The best peptide combination: LPQTLSR/AGR2, LYTYEPR/AGR3, SDLVNEEATGQFR/CEAM5, VTSLTACLVDQSLR/CD90, and GVCISPEAIVTDLPEDVK/SFRP4.

analyses indicated some of the biomarkers have potential utilities in the detection of prostate cancer.

Furthermore, our data show that peptides from the same protein with a good correlation produced similar AUC values. For example, the MMP9 peptides - AVIDDAFAR, FQTFEGDLK, LGLGADVAQVTGALR, SLGPALLLLQK - produced values of 0.82, 0.86, 0.86, and 0.86, respectively, as did the two well-correlated CD90 surrogate peptides: VLYLSAFTSK (0.86), HVLFGTVGVPEHTYR (0.87). VTSLTACLVDQSLR without significant correlations produced an AUC value of 0.77 (Table 2). The data suggests that the concentration of a given protein can be accurately quantified based on multiple well-correlated surrogate peptides. Multi-marker performance was also assessed by using multivariate analysis of various peptide combinations from different proteins (Figure 3 and Supplementary Table 7) and the combination of all surrogate peptides from the same protein (Supplementary Figure 3 and Supplementary Table 8). The best combination was LPQTLSR/AGR2, LYTYEPR/AGR3, SDLVNEEATGQFR/CEAM5, VTSLTACLVDQSLR/CD90, and

GVCISPEAIVTDLPEDVK/SFRP4 with $P = 0.002$ and AUC = 0.95.

Detection of clinically significant cancer by secreted protein markers

Next, we test the potential to differentiate high-risk cancer from low grade cancer. The prostate cancer cohort was grouped into either low volume/low grade (Gleason score ≤6 and tumor volume ≤ 0.5 cc [24]) or clinically significant (not meeting the above criteria for low volume/low grade disease, Supplementary Tables 9 and 10). The significance for most markers in identifying the high-risk cancers was not apparent except with MMP9 (Supplementary Table 10). The two surrogate peptides, FQTFEGDLK and LGLGADVAQVTGALR, produced P value of 0.022 in comparing low volume/low grade cancer and significant cancer (Figure 4 and Supplementary Table 10). The observation suggested an association between MMP9 and high grade/volume in this patient cohort. This result was also supported by cell-type transcriptomics data. Array signal intensity value for MMP9 in Gleason

Figure 4: Stratification of prostate cancer based on tumor volume and Gleason score. **(A)** The relative abundance ratios of FQTFEGDLK/MMP9 over IVGGWEC$_{cam}$EK/PSA between low volume/low grade cancer (n = 6) and significant cancer (n = 5), $P = 0.022$; **(B)** Urinary PSA concentrations (uPSA) between low volume/low grade cancer and significant cancer, $P = 0.93$; **(C)** Serum PSA concentrations (sPSA) between low volume/low grade cancer and significant cancer, $P = 0.32$.

4 cancer cells was 3004.10, ~12-fold higher than that of 238.41 in Gleason 3 cancer cells. (Supplementary Figure 2). For comparison, urine PSA and serum PSA showed no significance between the two cancer groups (Figure 4).

DISCUSSION

To date, disease detection relies mostly on single markers. The concept of multi-marker panel has the potential for more specific disease diagnosis and prognosis. Our data demonstrated the feasibility and promising aspects of multiplexed targeted MS assays for low-abundance prostate cancer-associated proteins in voided urine. The development for such assays is generally rapid in identifying the right surrogate peptides and implementation when compared with immunoassays that require the time-consuming generation of high quality monoclonal antibodies and their validation testing. With continuous advancement in measurement sensitivity, (e.g., LG-SRM [16] or PRISM-SRM [17]), SRM assays are feasible for sensitive measurement of low-abundance protein biomarkers in tissues [25, 26] and human body fluids [11, 27], as well as for facilitating the transition of biomarkers to large-scale clinical validation trials.

One important feature for targeted MS assays is that multiple surrogate peptides can be selected for a given protein. Each surrogate peptide from a given protein can serve as a unique marker since it may contain unique PTM or other sequence modifications. Without such modifications, the abundances for any two or more surrogate peptides from the same protein should correlate well across many samples. In studies involving human cell lines, most surrogate peptides (453/466) showed a high correlation coefficient ($R^2 > 0.8$) [21]. However, many surrogate peptides used in our urine analysis were found to have moderate correlation coefficients (median $R^2 = 0.70$) with data point deviations. This observation suggests that the target proteins in patient urine samples are more varied than those in single cell lines most likely due to allelic differences or isoforms. Therefore, multiple surrogate peptides per protein need to be tested in assay development and the individual peptide signatures may provide additional values for disease detection.

One challenge to urinary marker quantification is the large variation of urine protein concentration, and normalization strategies are often necessary. In our study, we observed that PSA as a prostate-specific secretory marker serves as an effective reference for normalization of other prostate cancer-associated proteins. Without PSA normalization, the performance for most markers was poor because of the multiple tissue sources of the urine proteome. Our assumption is that the main source of our panel of prostate cancer-associated proteins is from prostate cancer cells. By normalization against urinary PSA, a marker reflecting the total prostate cells, the marker performance was significantly improved. The significantly higher concentrations of urinary PSA found in some non-cancer samples could be due to donors with an enlarged prostate from benign hyperplasia. For example, prostate cancer patients with prostate volume of 35 cm³ (n = 29) and benign prostatic hyperplasia patients prostate volume of 45 cm³ (n = 35) were measured to have median urinary PSA levels of 52.6 ng/mL and 123.2 ng/mL, respectively [28].

The eventual goal of developing an informative panel of biomarkers is to reduce the need for prostate biopsy, an invasive, expensive, and potentially morbid procedure with up to a 4% risk of sepsis [29]. One could envision that prostate cancer diagnosis would involve the use of a relatively small number of markers as a tool for cancer detection, perhaps as a "reflex test" after PSA testing when the patient has an abnormal serum PSA. Notably, if the multi-marker panel is negative, no biopsy would be necessary especially when the negative predictive value is sufficiently high. Furthermore, our marker panel (e.g., MMP9) could have the potential utility in distinguishing low grade/low volume cancer from significant cancer. Therefore, by effectively integrating multi-marker measurement results, there is a greater possibility for detection of significant cancer with fewer biopsies performed in patients without cancer.

In conclusion, through comparison of cell type-specific transcriptomes, 14 cancer-associated secreted proteins were identified as candidate biomarkers. Sensitive multiplexed targeted MS assays were developed for reliable quantification of 10 secreted proteins (including previously reported AGR2) in human urine. All markers can be reproducibly detected and quantified in all the urine samples with at least one surrogate peptide. Most of the markers appear to be promising in prostate cancer detection in a pilot cohort study with initial AUC ranging from 0.75 to 0.86. Further studies with additional large sample cohorts to fully validate the performance of these markers are warranted. Our sensitive targeted SRM assays should also facilitate biomarker analysis of other cancers, especially for markers like secreted AGR2 that are widely present in many tumor types.

MATERIALS AND METHODS

Urine collection

The use of human urine samples was approved by the Institutional Review Boards of the University of Washington (UW), Pacific Northwest National Laboratory (PNNL), and the University of Texas Health Science Center at San Antonio (UTHSCSA). Samples from consented donors were anonymized before given to the researchers. Suffix N added to the sample codes denoted non-cancer, and suffix C denoted cancer from pre-op patients. Post-op urine was collected after surgical resection of the prostate.

Chemical reagents

Urea, dithiothreitol (DTT), iodoacetamide, ammonium formate, trifluoroacetic acid (TFA) and formic acid were purchased from Sigma (St. Louis, MO). The synthetic peptides labeled with $^{13}C/^{15}N$ on C-terminal lysine and arginine residues were from Thermo Scientific (San Jose, CA). The heavy peptides for PSA protein were estimated to be of >95% purity by HPLC.

Urine processing and protein digestion

Collected voided urine samples were processed within 2 h (to isolate RNA as well). The samples were centrifuged at 1,200 rpm for 5 min and the supernatant was stored at -80°C. Fifteen-90 mL of urine were desalted and concentrated using Amicon® Ultra-15 (3 kDa molecular weight cut-off, Millipore, Billerica, MA) [12]. Protein concentrations were determined by the BCA assay (Pierce, Rockford, IL). Concentrated urinary proteins from each sample, ranging from 200 to 300 µg, were denatured and reduced with 8 M urea and 10 mM DTT in 50 mM NH_4HCO_3, pH 8.0 for 1 h at 37°C. Protein cysteine residues were alkylated with 40 mM iodoacetamide for 1 h at room temperature in the dark. The resulting sample was diluted 6-fold with 50 mM NH_4HCO_3, pH 8.0, and digested by sequencing-grade modified porcine trypsin (Promega, Madison, WI) at 1:50 trypsin:protein (w/w) overnight at 37°C. The resulting digest was desalted by using 1 mL-SPE C18 column (Supelco, Bellefonte, PA) as described previously [11]. The final tryptic peptide concentration was determined by BCA. The peptide sample was diluted to 0.5 µg/µL with 0.1% formic acid in water, and crude heavy isotope-labeled synthetic peptides of protein markers were spiked in at an equimolar concentration of 10 fmol/µL; 10 fmol/µL of pure heavy peptide IVGGWEC$_{cam}$EK (C$_{cam}$: cysteine residue synthesized as carbamidomethyl cysteine) and 1 fmol/µL of pure heavy peptide LSEPAELTDAVK of PSA.

Database query

The human urine proteome databases archived in *PeptideAtlas* (http://www.peptideatlas.org) were queried for data entries of marker identifiers. The *UrinePA* build contained high confidence peptide and protein identifications obtained from five labs using tandem MS proteomics [30]. About 2,500 non-redundant proteins were cataloged at 1% false discovery rate. Another database listed 587 entries of a "Core Urinary Proteome", which was established from an in-depth analysis of second morning urine obtained over three days from seven healthy 25-35 year old volunteers [31].

SRM assays

Ten tryptic surrogate peptides were first chosen for the protein markers based on *in silico* trypsin digestion

and existing MS/MS data from our own lab, the Global Proteome Machine (GPM) and *PeptideAtlas*. These peptides were then evaluated by ESP predictor [32] and CONSeQuence [33] software. Three to five peptides with moderate hydrophobicity and high scores from the prediction tools were selected for peptide synthesis. The synthesized crude heavy-isotope labeled peptides were further evaluated in peptide response and fragmentation pattern. Optimal collision energy (CE) values were achieved by direct infusion of the individual peptides, and/or multiple LC-SRM runs with CE ramping. For each peptide, the three best transitions and matrix interference were determined. The relative intensity ratios among the three selected transitions for SRM were predefined by the internal standard heavy peptides in buffer. Matrix interference for a given transition that fell into mass widths Q1 and Q3 from co-eluting peptides was identified by a deviation from the expected relative intensity ratios among the transitions. The transition with no matrix interference was used for marker quantification in prostate urine samples. Before running the clinical cohort urine samples, the detectability of endogenous peptides in a pooled prostate cancer urine sample was systematically evaluated to finalize the best performing peptides for each protein marker. The detectable peptides were used for further quantification of the secreted protein markers in the cohort urine samples. For proteins with two or more detectable endogenous peptides, SRM signal correlation between any two surrogate peptides from the same protein was analyzed. For proteins with only one detectable endogenous peptide across all the urine samples, the potential of modifications on the surrogate peptides was evaluated by the knowledge-base information on PhosphoSitePlus and Uniport websites.

LG-SRM

The LG-SRM approach was previously demonstrated in enabling reproducible quantification of target proteins at ~10 ng/mL levels in nondepleted human serum [16]. Typically, 4 µL of tryptic digest samples with a peptide concentration of 0.5 µg/µL were directly loaded onto a capillary reversed-phase column, 75 µm inner diameter (i.d.) × 150 cm length, packed in-house with 3-µm Jupiter C18 bonded particles (Phenomenex, Torrance, CA) to permit long gradient separation without a trap column with its dead volume affecting peptide retention time. Peptide separations were performed at a mobile phase flow rate of 100 nL/min on a binary pump system using 0.1% formic acid in water as phase A and 0.1% formic acid in 90% acetonitrile as phase B. The profile for a 300 min gradient time was 5–15% B in 27 min, 15–25% B in 140 min, 25–35% B in 73 min, and 35–90% B in 60 min. The TSQ Vantage mass spectrometer was operated in the manner as previously described [16].

PRISM-SRM

The PRISM-SRM approach has been previously described for quantification of low-abundance proteins in human plasma or serum [17]. Briefly, high resolution reversed phase capillary LC with pH 10 mobile phase was used as the first dimensional separation of peptides from trypsin-digested human urine proteins. Following separation, the column eluent was automatically collected every minute into a 96-well plate during a ~100 min LC run while on-line SRM monitoring of heavy internal standard peptides was performed on a small split stream of the flow. Intelligent selection (termed *i*Selection) of target peptide fractions was achieved based on the on-line SRM signal of internal standard peptides. Prior to peptide fraction collection, 17 μL of water was added to each well to minimize excessive loss of peptides and to dilute the peptide fractions (~1:7) for LC-SRM analysis.

Following *i*Selection, the target peptide-containing fractions were subjected to LC-SRM measurement. All peptide fractions were analyzed by using the nanoACQUITY UPLC® system (Waters Corporation, Milford, MA) coupled on-line to a TSQ Vantage triple quadrupole mass spectrometer (Thermo Scientific, San Jose, CA). Solvents used were 0.1% formic acid in water (mobile phase A) and 0.1% formic acid in 90% acetonitrile (mobile phase B). Peptide separations were performed at a mobile phase flow rate of 400 nL/min using an ACQUITY UPLC BEH 1.7 μm C18 column (75 μm i.d. × 10 cm), which was connected to a chemically etched 20 μm i.d. fused-silica emitter via a Valco stainless steel union. Four μL of individual peptide fractions (total volume 20 μL) following PRISM were injected for LC separations using a binary gradient of 10-20% phase B in 7 min, 20-25% phase B in 17 min, 25-40% phase B in 1.5 min, 40-95% phase B in 2.5 min and 95% phase B in 6 min for a total time of ~35 min. The TSQ Vantage was operated in the same manner as previously described [11, 17]. A scan width of 0.002 *m/z* and a dwell time of 40 ms were set for all SRM transitions.

SRM data analysis

SRM data were analyzed using the Skyline software [34]. Peak detection and integration were determined based on (1) same retention time; (2) approximately same relative SRM peak intensity ratios across multiple transitions between light (L) peptides and heavy (H) peptide standards [11, 17, 35]. All data were manually inspected to ensure correct peak detection and accurate integration. Signal to noise ratio (S/N) was calculated by the peak apex intensity over the highest background noise in a retention time region of ±15 s for the target peptides [17, 35]. The background noise levels were conservatively estimated by visually inspecting chromatographic peak regions. Quantifiable endogenous surrogate peptides should have SRM signals with S/N ≥ 10. The RAW data from TSQ Vantage were loaded into Skyline to create high resolution figures of extracted ion chromatograms (XICs) of multiple transitions monitored for the target peptides = proteins.

Statistical analysis

GraphPad Prism (v.6.0) was used for statistical analysis and plotting; $P < 0.05$ was considered statistically significant [11]. Receiver operating characteristic (ROC) curves were produced in terms of sensitivity and specificity of protein markers at their specific cutoff values to evaluate the diagnostic performance of each candidate biomarker. The optimal cutoff was the point with the best sum of sensitivity and specificity. Multivariate evaluative analysis for various combinations of protein markers was done using SPSS (v.16.0) by logistic regression to find the best-fitting model for each comparison group.

Abbreviations

AGR2: anterior gradient 2; AUC: area under the receiver-operating characteristic curve; CE: collision energy; DRE: digital rectal exam; DTT: dithiothreitol; GPM: Global Proteome Machine; LC: liquid chromatography; LG: long gradient; L/H: the ratio of SRM signal from endogenous peptide over heavy-labeled internal standard; PRISM: high-pressure, high-resolution separation with intelligent selection and multiplexing; ROC: receiver operating characteristic; sPSA: serum PSA; SRM: selected reaction monitoring; TFA: trifluoroacetic acid; uPSA: urinary PSA; XIC: Extracted ion chromatogram.

ACKNOWLEDGMENTS

REFERENCES

1. Hernandez J, Thompson IM. Prostate-specific antigen: a review of the validation of the most commonly used cancer biomarker. Cancer. 2004; 101:894-904.

2. Esserman LJ, Thompson IM, Jr. and Reid B. Overdiagnosis and overtreatment in cancer: an opportunity for improvement. JAMA. 2013; 310:797-798.

3. Sokoll LJ, Ellis W, Lange P, Noteboom J, Elliott DJ, Deras IL, Blase A, Koo S, Sarno M, Rittenhouse H, Groskopf J, Vessella RL. A multicenter evaluation of the PCA3 molecular urine test: pre-analytical effects, analytical performance, and diagnostic accuracy. Clin Chim Acta. 2008; 389:1-6.

4. van Gils MP, Hessels D, van Hooij O, Jannink SA, Peelen WP, Hanssen SL, Witjes JA, Cornel EB, Karthaus HF, Smits GA, Dijkman GA, Mulders PF, Schalken JA. The time-resolved fluorescence-based PCA3 test on urinary sediments after digital rectal examination; a Dutch multicenter validation of the diagnostic performance. Clin Cancer Res. 2007; 13:939-943.

5. Hessels D, van Gils MP, van Hooij O, Jannink SA, Witjes JA, Verhaegh GW, Schalken JA. Predictive Value of PCA3 in urinary sediments in determining clinico-pathological characteristics of prostate cancer. Prostate. 2010; 70:10-16.

6. Durand X, Xylinas E, Radulescu C, Haus-Cheymol R, Moutereau S, Ploussard G, Forgues A, Robert G, Vacherot F, Loric S, Allory Y, Ruffion A and de la Taille A. The value of urinary prostate cancer gene 3 (PCA3) scores in predicting pathological features at radical prostatectomy. Bju Int. 2012; 110:43-49.

7. Day JR, Jost M, Reynolds MA, Groskopf J, Rittenhouse H. PCA3: From basic molecular science to the clinical lab. Cancer Lett. 2011; 301:1-6.

8. Pascal LE, Goo YA, Vencio RZ, Page LS, Chambers AA, Liebeskind ES, Takayama TK, True LD, Liu AY. Gene expression down-regulation in CD90(+) prostate tumor-associated stromal cells involves potential organ-specific genes. BMC Cancer. 2009: 9:317.

9. Pascal LE, Vencio RZ, Page LS, Liebeskind ES, Shadle CP, Troisch P, Marzolf B, True LD, Hood LE, Liu AY. Gene expression relationship between prostate cancer cells of Gleason 3, 4 and normal epithelial cells as revealed by cell type-specific transcriptomes. BMC Cancer. 2009; 9:452.

10. Wayner EA, Quek SI, Ahmad R, Ho ME, Loprieno MA, Zhou Y, Ellis WJ, True LD, Liu AY. Development of an ELISA to detect the secreted prostate cancer biomarker AGR2 in voided urine. Prostate. 2012; 72:1023-1034.

11. Shi T, Gao Y, Quek SI, Fillmore TL, Nicora CD, Su D, Zhao R, Kagan J, Srivastava S, Rodland KD, Liu T, Smith RD, Chan DW, Camp DG, 2nd, Liu AY, Qian WJ. A highly sensitive targeted mass spectrometric assay for quantification of AGR2 protein in human urine and serum. J Proteome Res. 2014; 13:875-882.

12. True LD, Zhang H, Ye ML, Huang CY, Nelson PS, von Haller PD, Tjoelker LW, Kim JS, Qian WJ, Smith RD, Ellis WJ, Liebeskind ES, Liu AY. CD90/THY1 is overexpressed in prostate cancer-associated fibroblasts and could serve as a cancer biomarker. Modern Pathol. 2010; 23:1346-1356.

13. Bu H, Bormann S, Schafer G, Horninger W, Massoner P, Neeb A, Lakshmanan VK, Maddalo D, Nestl A, Sultmann H, Cato AC, Klocker H. The anterior gradient 2 (AGR2) gene is overexpressed in prostate cancer and may be useful as a urine sediment marker for prostate cancer detection. Prostate. 2011; 71:575-587.

14. Shi T, Su D, Liu T, Tang K, Camp DG, 2nd, Qian WJ, Smith RD. Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics. Proteomics. 2012; 12:1074-1092.

15. Shi T, Song E, Nie S, Rodland KD, Liu T, Qian WJ, Smith RD. Advances in targeted proteomics and applications to biomedical research. Proteomics. 2016; 16:2160-2182.

16. Shi T, Fillmore TL, Gao Y, Zhao R, He J, Schepmoes AA, Nicora CD, Wu C, Chambers JL, Moore RJ, Kagan J, Srivastava S, Liu AY, Rodland KD, Liu T, Camp DG, 2nd, Smith RD, Qian WJ. Long-gradient separations coupled with selected reaction monitoring for highly sensitive, large scale targeted protein quantification in a single analysis. Anal Chem. 2013; 85:9196-9203.

17. Shi T, Fillmore TL, Sun X, Zhao R, Schepmoes AA, Hossain M, Xie F, Wu S, Kim JS, Jones N, Moore RJ, Pasa-Tolic L, Kagan J, Rodland KD, Liu T, Tang K. Antibody-free, targeted mass-spectrometric approach for quantification of proteins at low picogram per milliliter levels in human plasma/serum. Proc Natl Acad Sci U S A. 2012; 109:15395-15400.

18. Shi T, Qian WJ. Antibody-free PRISM-SRM for multiplexed protein quantification: is this the new competition for immunoassays in bioanalysis? Bioanalysis. 2013; 5:267-269.

19. Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. Mol Cell Proteomics. 2007; 6:2212-2229.

20. Nie S, Shi T, Fillmore TL, Schepmoes AA, Brewer H, Gao Y, Song E, Wang H, Rodland KD, Qian WJ, Smith RD, Liu T. Deep-dive targeted quantification for ultrasensitive analysis of proteins in nondepleted human blood plasma/serum and tissues. Anal Chem. 2017; 89:9139-9146.

21. Worboys JD, Sinclair J, Yuan Y, Jorgensen C. Systematic evaluation of quantotypic peptides for targeted analysis of the human kinome. Nat Methods. 2014; 11:1041-1044.

22. Zhang Q, Fillmore TL, Schepmoes AA, Clauss TR, Gritsenko MA, Mueller PW, Rewers M, Atkinson MA, Smith RD, Metz TO. Serum proteomics reveals systemic dysregulation of innate immunity in type 1 diabetes. J Exp Med. 2013; 210:191-203.

23. Marks LS, Bostwick DG. Prostate cancer specificity of PCA3 gene testing: examples from clinical practice. Rev Urol. 2008; 10:175-181.

24. Nakanishi H, Groskopf J, Fritsche HA, Bhadkamkar V, Blase A, Kumar SV, Davis JW, Troncoso P, Rittenhouse H, Babaian RJ. PCA3 molecular urine assay correlates with prostate cancer tumor volume: implication in selecting candidates for active surveillance. J Urol. 2008; 179:1804-1809; discussion 1809-1810.

25. He J, Sun X, Shi T, Schepmoes AA, Fillmore TL, Petyuk VA, Xie F, Zhao R, Gritsenko MA, Yang F, Kitabayashi N, Chae SS, Rubin MA, Siddiqui J. Antibody-independent targeted quantification of TMPRSS2-ERG fusion protein products in prostate cancer. Mol Oncol. 2014; 8:1169-1180.

26. He J, Schepmoes AA, Shi T, Wu C, Fillmore TL, Gao Y, Smith RD, Qian WJ, Rodland KD, Liu T, Camp DG, 2nd, Rastogi A, Tan SH. Analytical platform evaluation for quantification of ERG in prostate cancer using protein and mRNA detection methods. J Transl Med. 2015; 13:54.

27. Engholm-Keller K, Larsen MR. Technologies and challenges in large-scale phosphoproteomics. Proteomics. 2013; 13:910-931.

28. Bolduc S, Lacombe L, Naud A, Gregoire M, Fradet Y, Tremblay RR. Urinary PSA: a potential useful marker when serum PSA is between 2.5 ng/mL and 10 ng/mL. Can Urol Assoc. 2007; 1:377-381.

29. Nam RK, Saskin R, Lee Y, Liu Y, Law C, Klotz LH, Loblaw DA, Trachtenberg J, Stanimirovic A, Simor AE, Seth A, Urbach DR, Narod SA. Increasing hospital admission rates for urological complications after transrectal ultrasound guided prostate biopsy. J Urol. 2013; 189:S12-17; discussion S17-18.

30. Farrah T, Deutsch EW, Omenn GS, Sun Z, Watts JD, Yamamoto T, Shteynberg D, Harris MM, Moritz RL. State of the human proteome in 2013 as viewed through PeptideAtlas: comparing the kidney, urine, and plasma proteomes for the biology- and disease-driven Human Proteome Project. J Proteome Res. 2014; 13:60-75.

31. Nagaraj N, Mann M. Quantitative analysis of the intra- and inter-individual variability of the normal urinary proteome. J Proteome Res. 2011; 10:637-645.

32. Fusaro VA, Mani DR, Mesirov JP, Carr SA. Prediction of high-responding peptides for targeted protein assays by mass spectrometry. Nat Biotechnol. 2009; 27:190-198.

33. Eyers CE, Lawless C, Wedge DC, Lau KW, Gaskell SJ, Hubbard SJ. CONSeQuence: prediction of reference peptides for absolute quantitative proteomics using consensus machine learning approaches. Mol Cell Proteomics. 2011; 10:M110.003384.

34. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 2010; 26:966-968.

35. Shi T, Sun X, Gao Y, Fillmore TL, Schepmoes AA, Zhao R, He J, Moore RJ, Kagan J, Rodland KD, Liu T, Liu AY, Smith RD, Tang K, Camp DG 2nd, Qian WJ. Targeted quantification of low ng/mL level proteins in human serum without immunoaffinity depletion. J Proteome Res. 2013; 12:3353-3361.

PANEL OF PROSTATE CANCER-ASSOCIATED SECRETED PROTEINS FOR DIAGNOSIS AND PROGNOSIS

FIELD

5   This application provides methods of diagnosing and/or treating prostate cancer, following determining expression levels of several prostate cancer-related secretory molecules, AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4, for example in a urine sample.

ACKNOWLEDGMENT OF GOVERNMENT SUPPORT

PARTIES TO JOINT RESEARCH AGREEMENT

20   This application describes and claims certain subject matter that was developed under a written joint research agreement between Battelle Memorial Institute and the University of Washington.

BACKGROUND

25   Prostate cancer is the one of the most common cancers diagnosed in the United States; it is only second to skin cancer in men (Matrana et al., J. Adv. Pract. Oncol., 4(1):16-21, 2013). It is also the second most common cause of cancer death in men in the U.S., wherein approximately 26,120 men will die of prostate cancer in 2016 (Matrana et al., J. Adv. Pract. Oncol., 4(1):16-21, 2013; American Cancer Society,

30   *Cancer Facts & Figures 2016*, Atlanta: American Cancer Society, 2016). Further, a

prostate cancer diagnosis can vary from a low- or intermediate-risk to a high-risk diagnosis, which means that the patient has a low-, intermediate-, or high-risk of pathological and biochemical outcomes after treatment (*e.g.*, by radical prostatectomy); metastasis; prostate cancer-specific mortality; and all-cause mortality (Cooperberg *et al.*, J Cancer Inst., 101(12):878-887, 2009). A diagnosis of a particular risk category can determine the course of treatment, which can, for example, range dramatically from monitoring, in the case of low-risk cancers, to radiation or surgical procedures for higher risk cancers (Bill-Axelson et al., N. Engl. J. Med., 370:932-942, 2014).

The common methods of screening men for prostate cancer are a digital rectal exam (DRE) and serum prostate-specific antigen (sPSA) screening, both of which can be administered alone or in combination (Zhang et al., UWOMJ, 82(1):10-13, 2013). The use of DRE is not favored by patients, as it can be an invasive and uncomfortable process (Zhang et al., UWOMJ, 82(1):10-13, 2013). Further, sPSA testing is also controversial (Matrana et al., J. Adv. Pract. Oncol., 4(1):16-21, 2013). The advent of screening for prostate cancer using sPSA testing has decreased the number of cases diagnosed with metastasis and, thus, prostate cancer mortality (Matrana et al., J. Adv. Pract. Oncol., 4(1):16-21, 2013). However, the medical community remains divided over its overall utility, as studies have raised concerns over the potential negative effects associated with sPSA testing, including unnecessary biopsies, overdiagnosis, and/or overtreatment of patients (Matrana et al., J. Adv. Pract. Oncol., 4(1):16-21, 2013).

Human urine is an ideal clinical specimen for testing prostate cancer biomarkers as prostatic secretion passes into the urine. Urine is a natural waste product; its collection is the least invasive and poses no health risk to the patients. Furthermore, urine can be conveniently collected multiple times compared with human blood. Currently, one prostate cancer urine test measures the cancer-specific non-coding transcript PCA3 from released prostate cancer cells [1]. This PCA3 test involves target capture and amplification with chemiluminescent probe detection. In a cohort of >500 patients with sPSA between 3 and 15 ng/mL, the area under the receiver-operating characteristic curve (AUC) was 0.66 with a sensitivity of 65% and a specificity of 66% [2]. As a prognostic marker, PCA3 showed no significant link to Gleason score, tumor

volume, and cancer stage in a cohort of 70 cases [3], though a link to tumor volume and surgical margin was reported in another study [4]. PCA3 is a low abundance transcript, and an "attentive" DRE by an experienced urologist is required to enhance the PCA3 signal [5].

5   Thus, while screening men for prostate cancer decreases mortality from the disease, a need remains in the medical community for a more accurate and non-invasive means of screening for prostate cancer.

SUMMARY

10   It is shown herein that a combination of prostate cancer-related molecules, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4, can be detected in patient urine samples (such as urine from a human subject) and can be used as biomarkers to diagnose and/or treat a patient with or at risk for prostate cancer.

15   There is a need for accurate and non-invasive means of screening for prostate cancer. For example, common means of screening for prostate cancer include a digital rectal exam (DRE), which is considered uncomfortable and invasive, and a serum prostate-specific antigen (sPSA) test, which is sufficiently inaccurate to raise concerns about overtreating and overdiagnosing subjects at risk for prostate cancer.

20   Moreover, there is also a need for an accurate, non-invasive means of distinguishing prostate cancer in different risk categories with the concern of overtreatment. For example, low-risk prostate cancer patients can remain healthy for a long time without undergoing invasive and painful surgery with potentially harmful side effects, and, thus, are considered ideal candidates for observation-based therapies, such

25   as watchful waiting and active surveillance.

          The inventors have identified a panel of prostate-secreted biomarkers and biomarker combinations detectable in urine (*e.g.*, AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4 as well as the biomarker combinations listed in FIG. 15) that provide an accurate assessment of

30   whether or not a subject is likely to have prostate cancer. Further, the inventors have

also identified a biomarker (MMP9) that can distinguish with significant accuracy subjects with low-risk prostate cancer from subjects with intermediate- or high-risk prostate cancer.

Methods are provided for treating a subject with prostate cancer. Such methods can include measuring expression of at least two prostate cancer-related molecules in a sample obtained from a subject, including the prostate cancer-related molecules AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4. In some examples, expression levels are normalized to PSA levels detected in urine. The methods can further include measuring increased expression of the at least two prostate cancer-related molecules, for example as compared to a control or reference value representing expression for each of the at least two prostate cancer-related molecules expected in a sample from a subject who does not have prostate cancer. In addition, the methods can include administering at least one of watchful waiting, active surveillance, surgery, radiation, hormone therapy, chemotherapy, brachytherapy, cryotherapy, ultrasound, bisphosphate therapy, biologic therapy, or vaccine therapy to the subject with prostate cancer, thereby treating the subject.

Methods are provided for diagnosing prostate cancer in a subject. The methods can include detecting the expression of at least two prostate cancer-related molecules in a sample obtained from a subject, including the prostate cancer-related molecules comprising AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4. In some examples, expression levels are normalized to PSA levels detected in urine. The methods can further include comparing the expression of the at least two prostate cancer-related molecules in the sample obtained from the subject to at least one control or reference value representing expression for each of the at least two prostate cancer-related molecules expected in a sample from a subject who does not have prostate cancer. In addition, the methods can include determining that the subject has prostate cancer when increased expression of the at least two prostate cancer-related molecules between the sample and the control is detected.

In some examples, the at least two prostate cancer-related molecules can include all of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4; the combinations of prostate cancer-related molecules listed in FIG. 15; low-abundance molecules, such as AGR2, AGR3, CCL3,

5   CEACAM5, and CEACAM6; and moderate-to-low-abundance molecules, such as CRISP3, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4.

Methods are provided for treating a subject with intermediate- or high-risk prostate cancer. The methods can include measuring expression of MMP9 in a sample obtained from a subject, which can be determined based on MMP9 protein

10   concentration using at least one surrogate peptide of MMP9, and the surrogate peptide can be at least one of FQTFEGDLK or LGLGADVAQVTGALR. In some examples, the MMP9 expression level is normalized to PSA levels detected in urine. The methods further include measuring increased expression of MMP9 in the sample obtained from the subject as compared to a control or reference value representing expression of

15   MMP9 expected in a sample from a subject who has low-risk prostate cancer. In addition, the methods can include administering treatment for intermediate- or high-risk prostate cancer, thereby treating the subject.

Methods are provided for diagnosing intermediate- or high-risk prostate cancer. The methods can include detecting expression of MMP9 in a sample obtained from a

20   subject, which can be determined based on MMP9 protein concentration using at least one surrogate peptide of MMP9, and the surrogate peptide can be at least one of FQTFEGDLK or LGLGADVAQVTGALR. In some examples, expression levels are normalized to PSA levels detected in urine. In further examples, the methods can include comparing the expression of MMP9 in the sample obtained from the subject to

25   expression of MMP9 expected in a sample from a subject who has low-risk prostate cancer (*e.g.*, a reference value representing MMP9 expression expected in a subject with low-risk prostate cancer). In addition, the methods can include determining that the subject has intermediate- or high-risk prostate cancer when increased expression of MMP9 between the sample and the control is detected.

In some examples, the methods can include determining expression based on protein concentration, which can be determined using the concentration of at least one surrogate peptide of the protein, such as a peptide listed in FIG. 4. In other examples, the methods can include determining the protein concentration using an immunoassay, such as an ELISA. In additional examples, the protein concentration can determined using mass spectrometry, for example, using LC-SRM, LG-SRM, or PRISM-SRM.

In some other examples, the methods include normalizing expression of the at least two prostate cancer-related molecules to the amount of a prostate protein, such as PSA, which can be determined using, for example, at least one surrogate peptide of PSA, such as IVGGWECEK or LSEPAELTDAVK.

In additional examples, the sample can be a urine sample, and, in further examples, the subject can be a human subject. In particular examples, the expression detected can have an AUC value of greater than 0.8, such as at least 0.85, at least 0.9, or at least 0.95. In particular examples, the disclosed methods have a sensitivity of at least 78 to 86% and specificity of at least 100%.

The foregoing and other objects and feature of the disclosure will become more apparent from the following detailed description, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows a panel that lists the query result from the database UrinePA. The parameter "observed" indicates abundance.

FIG. 2 shows thirteen prostate cancer-associated secreted proteins and their surrogate peptides. For each surrogate peptide, the three best transitions without co-eluting interference were monitored. The "a" indicates that endogenous light peptides were detected in the pooled urine sample. The "b" indicates that cysteine was synthesized as carbamidomethyl cysteine.

FIGS. 3A-3B show extracted ion chromatograms (XICs) of detected proteins in a single urine sample, P07-031C. Seven proteins (CD90, CRISP3, CXCL14, IL24, MMP9, POSTN, and SFRP4) were detected by LG-SRM, and the other five (AGR2,

AGR3, CCL3, CEACAM5, and CEACAM6) were in extremely low abundance and were detected by PRISM-SRM. The monitored transitions for the surrogate peptides of each protein are THY1/CD90: VLYLSAFTSK, 564.8/916.5 (blue), 564.8/640.3 (chestnut), and 564.8/753.4 (purple); CRISP3: WANQCcamNYR, 556.2/854.3 (purple),

5   556.2/925.4 (blue), and 556.2/612.3 (chestnut); CXCL14: MVIITTK, 403.2/575.4 (purple), 403.2/674.4 (blue), and 403.2/462.3 (chestnut); IL24: LWEAFWAVK, 575.3/850.4 (blue), 575.3/721.4 (purple), and 575.3/650.4 (chestnut); MMP9: AVIDDAFAR, 489.3/807.4 (blue), 489.3/694.3 (purple), and 489.3/579.3 (chestnut); POSTN: AAAITSDILEALGR, 700.9/1074.8 (blue), 700.9/973.5 (purple), and

10   700.9/771.5 (chestnut); SFRP4: GVCcamISPEAIVTDLPEDVK, 971.5/587.3 (chestnut), 971.5/916.5 (blue), and 971.5/1425.7 (purple); AGR2: LPQTLSR, 407.7/351.2 (chestnut), 407.7/476.2 (purple), and 407.7/604.3 (blue); AGR3: LYTYEPR, 471.2/665.3 (blue), 471.2/272.2 (purple), and 471.2/277.2 (chestnut); CCL3: QVCcamADPSEEWVQK, 788.4/1002.5 (chestnut), 788.4/1117.5 (purple), and

15   788.4/1188.6 (blue); CEACAM5: SDLVNEEATGQFR, 733.3/679.4 (blue), 733.3/1051.5 (chestnut), and 733.3/937.4 (purple); CEACAM6: SDPVTLNVLYGPDGPTISPSK, 1079.1/1055.5 (blue), 1079.1/331.2 (chestnut), and 1079.1/998.5 (purple).

          FIG. 4 shows prostate cancer-associated secreted proteins and their surrogate

20   peptides. The "a" indicates that these surrogate peptides were confidently detected in the pooled urine sample. The "b" indicates that the cysteine was synthesized as carbamidomethyl cysteine.

          FIGS. 5A-5F show correlation plots between any two MMP9 surrogate peptides in 20 urine samples. **FIG. 5A** shows a relative abundance correlation between

25   FQTFEGDLK (*y*-axis) and AVIDDAFAR (*x*-axis); **FIG. 5B** shows a relative abundance correlation between LGLGADVAQVTGALR (*y*-axis) and AVIDDAFAR (*x*-axis); **FIG. 5C** shows a relative abundance correlation between SLGPALLLLQK and AVIDDAFAR; **FIG. 5D** shows a relative abundance correlation between SLGPALLLLQK (*y*-axis) and LGLGADVAQVTGALR (*x*-axis); **FIG. 5E** shows a

30   relative abundance correlation between LGLGADVAQVTGALR (*y*-axis) and

FQTFEGDLK (*x*-axis); **FIG. 5F** shows a relative abundance correlation between SLGPALLLLQK (*y*-axis) and FQTFEGDLK (*y*-axis). L/H = the ratio of SRM signal from endogenous peptide over heavy-labeled internal standard. R^2 values range from 0.59 to 0.93.

5   **FIGS. 6A-6B** show correlation curves between two surrogate peptides from the same protein. **FIG. 6A** shows CD90 (with the removal of the point with the red circle, the correlation coefficient of R^2 significantly decreased from 0.70 to 0.21); **FIG. 6B** shows CRISP3 (with the removal of the point with the red circle, the correlation coefficient of R^2 significantly decreased from 0.65 to 0.14).

10   **FIGS. 7A-7B** show correlation curves between two PSA surrogate peptides in 20 urine subjects. **FIG. 7A** shows the entire range of relative abundance of the peptides; **FIG. 7B** shows a small range of relative abundance. Data points with the blue dash circle significantly deviated from the correlation curve.

          FIG. 8 shows a summary of SRM measurements of PSA protein in 27 clinical
15   urine samples including 7 post-op subjects (two purified PSA internal standards, IVGGWECcamEK and LSEPAELTDAVK, were spiked at 1 fmol/µL and 10 fmol/ µL, respectively). The "a" indicates that the L/H ratios were corrected according to the correlation curve of the two PSA surrogate peptides as well as the experimental observation (P06017Pre: the measured value of 0.828 was changed into 1.656;
20   P07040Pre: the measured value of 0.261 was changed into 0.366; P08015Pre: the measured value of 3.240 was changed into 4.739).

          FIG. 9 shows extracted ion chromatograms of transitions monitored for the PSA peptide LSEPAE(L/I)TDAVK in urine from two patients (for the two deviated data points on the correlation curve at the low concentration range in **FIG. 7B**).

25   **FIG. 10** shows an estimation of the percentage of PSA from the post-op urine over PSA from the non-cancer urine (the surrogate peptide IVGGWEC$_{cam}$EK was used). The "a" indicates that the L/H ratios were corrected according to the correlation curve of the two PSA surrogate peptides (P07040Pre: the measured value of 0.261 was changed into 0.366; P08015 Pre: the measured value of 3.240 was changed into 4.739).

FIGS. 11A-11B show a summary of multiplex SRM measurements of prostate cancer-associated secreted proteins in 20 clinical urine samples.

FIG. 12 shows the performance of surrogate peptide markers derived from 10 prostate cancer-associated secreted proteins in 20 urine samples (14 cancer and 6 non-cancer samples). The "a" indicates that P values were obtained from the Mann-Whitney U test. The "b" indicates that these are the sensitivity and specificity at the optimal cutoff point (i.e., the best sum of the sensitivity and specificity). The "c" indicates that the cysteine was synthesized as carbamidomethyl cysteine.

FIGS. 13A-13B show the SRM signal ratio of urinary secreted protein/PSA, i.e., $(L/H)_{peptide\ marker}/(L/H)_{PSA}$ from SRM measurements in 20 clinical urine samples (crude internal standards for prostate cancer-associated secreted proteins and purified internal standard for PSA surrogate peptide IVGGWECcamEK were spiked at 10 fmol/µL and 1 fmol/µL, respectively). The "a" indicates that the L/H ratios were corrected according to the correlation curve of the two PSA surrogate peptides (P07040Pre: the measured value of 0.261 was changed into 0.366; P08015Pre: the measured value of 3.240 was changed into 4.739).

FIGS. 14A-14F show urine protein biomarkers for prostate cancer. **FIG. 14A** shows the CEACAM5 relative abundance between non-cancer (n = 6) and cancer urine (n = 14; *P* = 0.322); **FIG. 14B** shows an ROC curve analysis of the relative abundance of CEACAM5 in the measured 20 urine samples; **FIG. 14C** shows CEACAM5/PSA concentration ratios between non-cancer and cancer (*P* = 0.012); **FIG. 14D** shows an ROC curve analysis of the CEACAM5/PSA concentration ratios; **FIG. 14E** shows significant differentiation between non-cancer and cancer (*P* = 0.0034) with the use of the best peptide combination; **FIG. 14F** shows an ROC curve analysis of the best peptide combination. The relative abundances of CEACAM5 and PSA were derived from their surrogate peptides, SDLVNEEATGQFR and IVGGWECcamEK, respectively. The best peptide combination: LPQTLSR/AGR2, LYTYEPR/AGR3, SDLVNEEATGQFR/CEACAM5, VTSLTACLVDQSLR/CD90, and GVCISPEAIVTDLPEDVK/SFRP4.

FIGS. 15 shows selected combinations of multiple markers for achieving better discrimination than individual markers between cancer and non-cancer. The "a" indicates that the P values were obtained from the Mann-Whitney U test. The "b" indicates that these are the sensitivity and specificity at the optimal cutoff point (i.e., the

5   best sum of the sensitivity and specificity). The "c" indicates that the cysteine was synthesized as carbamidomethyl cysteine.

FIGS. 16A-16B show urinary AGR2 measured by ELISA. **FIG. 16A** shows an ROC plot for urinary AGR2 determined by ELISA. The P value obtained for this cohort was 0.01. **FIG. 16B** shows the urinary levels of AGR2 in a non-cancer healthy male

10   (P16-050A-J), which were measured by ELISA using samples donated within a period of 14 days. Buffer was the negative control, and PC3 was the positive control. In bar 4, alcohol was added to P16-050A urine before the ELISA was performed.

FIG. 17 shows PSA concentrations in urine and serum for 20 measured subjects (urinary PSA and serum PSA concentrations were obtained from SRM measurements

15   and ELISA measurements, respectively). The "a" indicates that the L/H ratios were corrected according to the correlation curve of the two PSA surrogate peptides (P07040Pre: the measured value of 0.261 was changed into 0.366; P08015Pre: the measured value of 3.240 was changed into 4.739).

FIGS. 18A-18B show the ratios of secreted protein over PSA concentrations,

20   urinary PSA (uPSA) and serum PSA between low volume/low grade cancer (n = 6) and significant cancer (n = 5). The low volume/low grade cancer: Gleason score ≤ 6 and tumor volume ≤ 0.5 cc; the significant cancer: Gleason score > 6 and tumor volume > 0.5 cc.

FIGS. 19A-19C show stratification of prostate cancer based on tumor volume

25   and Gleason score. **FIG. 19A** shows the relative abundance ratios of FQTFEGDLK/MMP9 over IVGGWECcamEK/PSA between low volume/low grade cancer (n = 6) and significant cancer (n = 5), P = 0.022; **FIG. 19B** shows urinary PSA concentrations (uPSA) between low volume/low grade cancer and significant cancer (P = 0.93); **FIG. 19C** shows a comparison of serum PSA concentrations (sPSA) between

30   low volume/low grade cancer and significant cancer (P = 0.32).

FIGS. 20A-20B show gene expression levels of protein markers in Gleason 3 (labeled 05-179_CD26t) vs. Gleason 4 (labeled 08-032_CP_epi_CD26posi) cancer cells. Differential expression is displayed using gray scale (**FIG. 20A**) and histogram (**FIG. 20B**) formats.

5

DETAILED DESCRIPTION

The following explanations of terms and methods are provided to better describe the present disclosure and to guide those of ordinary skill in the art in the practice of the present disclosure. The singular forms "a," "an," and "the" refer to one or more than

10   one, unless the context clearly dictates otherwise. For example, the term "comprising a protein" includes single or plural cells and is considered equivalent to the phrase "comprising at least one protein." The term "or" refers to a single element of stated alternative elements or a combination of two or more elements, unless the context clearly indicates otherwise. As used herein, "comprises" means "includes." Thus,

15   "comprising A or B," means "including A, B, or A and B," without excluding additional elements. Dates of GenBank® Accession Nos. referred to herein are the sequences available at least as early as December 12, 2016. All references and GenBank® Accession numbers cited herein are incorporated by reference.

Unless explained otherwise, all technical and scientific terms used herein have

20   the same meaning as commonly understood to one of ordinary skill in the art to which this disclosure belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are described below. The materials, methods, and examples are illustrative only and not intended to be limiting.

25   In order to facilitate review of the various embodiments of the disclosure, the following explanations of specific terms are provided.

Administration: To provide or give a subject a therapeutic intervention, such as a therapeutic drug, procedure, or protocol (*e.g.*, for a subject with prostate cancer, docetaxel, prostatectomy, and active surveillance, respectively). Exemplary routes of

30   administration for drug therapy include, but are not limited to, oral, injection (such as

subcutaneous, intramuscular, intradermal, intraperitoneal, intratumoral, intraprostatic, and intravenous), sublingual, rectal, transdermal, intranasal, and inhalation routes.

Anterior gradient 2 (AGR2): Also known as AG2 (*e.g.*, OMIM 606358); protein disulfide isomerase family A, member 17 (PD1A17 or member 17); or secreted

5   cement gland protein XAG-2 homolog, AGR2 belongs to the protein disulfide isomerase (PDI) family. The strongest AGR2 expression has been measured in the lung, pancreas, trachea, stomach, colon, prostate, and small intestine as well as in certain breast cancer cell lines. AGR2 plays a role in regulating the response to DNA damage as well as cell migration, growth, proliferation, and transformation. AGR2

10   overexpression plays a role in cancer and metastasis.

Includes AGR2 nucleic acid molecules and proteins. AGR2 sequences are publicly available. For example, GenBank® Accession Nos. NM_006408.3, NM_001106725.1, and NM_011783.2 disclose exemplary human, rat, and mouse AGR2 nucleotide sequences, respectively, and GenBank® Accession Nos.

15   NP_006399.1, NP_001100195.1, and NP_035913.1 disclose exemplary human, rat, and mouse AGR2 protein sequences, respectively. One of ordinary skill in the art can identify additional AGR2 nucleic acid and protein sequences, including AGR2 variants that retain AGR2 biological activity (such as having increased levels in urine from a subject with prostate cancer).

20   **Anterior gradient 3 (AGR3)**: Also known as AG3 (*e.g.*, OMIM 609482); protein disulfide isomerase family A, member 18 (PD1A18 or member 18); or breast cancer membrane protein 11 (BCMP11), AGR3 belongs to the protein disulfide isomerase (PDI) family. AGR3 is expressed in certain epithelial and cancerous cells, such as breast and prostate cancer cells as well as cancerous epithelial cells, but exhibits

25   restricted expression in most normal cells. AGR3 expression is associated with ciliary function in epithelial cells as well as cell differentiation and proliferation in ovarian and breast cancers.

Includes AGR3 nucleic acid molecules and proteins. AGR3 sequences are publicly available. For example, GenBank® Accession Nos. NM_176813.4,

30   NM_001106724.1, and NM_207531.3, disclose exemplary human, rat, and mouse

AGR3 nucleotide sequences, respectively, and GenBank® Accession Nos. NP_789783.1, NP_001100194.1, and NP_997414.2 disclose exemplary human, rat, and mouse AGR3 protein sequences, respectively. One of ordinary skill in the art can identify additional AGR3 nucleic acid and protein sequences, including AGR3 variants

5   that retain AGR3 biological activity (such as having increased levels in urine from a subject with prostate cancer).

Cluster Designation 90 (CD90): Also known as THY-1 T-cell antigen (THY1; *e.g.*, OMIM 188230), CD90 is a cell surface glycoprotein in the immunoglobulin superfamily. CD90 is expressed by multiple cell types, including endothelial, smooth

10   muscle, bone marrow, umbilical cord blood, fibroblasts and hemopoietic cells, and in tissues such as nervous and lymphoid tissues. Further, CD90 functions as a tumor suppressor and plays a role in cell adhesion and communication as well as immunity.

Includes CD90 nucleic acid molecules and proteins. CD90 sequences are publicly available. For example, GenBank® Accession Nos. BC065559.1,

15   NM_012673.2, and NM_009382.3 disclose exemplary human, rat, and mouse CD90 nucleotide sequences, respectively, and GenBank® Accession Nos. NP_001298091.1, NP_036805.1, and NP_033408.1 disclose exemplary human, rat, and mouse CD90 protein sequences, respectively. One of ordinary skill in the art can identify additional CD90 nucleic acid and protein sequences, including CD90 variants that retain CD90

20   biological activity (such as having increased levels in urine from a subject with prostate cancer).

Carcinoembryonic antigen-related cell adhesion molecule 5 (CEACAM5): Also known as cluster of differentiation 66e (CD66e; *e.g.*, OMIM 114890), CEACAM5 is an immunoreactive glycoprotein in the CEA (carcinoembryonic antigen) family.

25   Many CEACAM family proteins are expressed in hematopoietic cells, and CEACAM5 plays a role in cell signaling, adhesion, differentiation, apoptosis, and polarity. Further, elevated levels of CEA proteins have been found in colorectal and other cancers as well as in patients with benign liver disease.

Includes CEACAM5 nucleic acid molecules and proteins. CEACAM5

30   sequences are publicly available. For example, GenBank® Accession Nos.

NM_004363.5 and NM_028480.2 disclose exemplary human and mouse CEACAM5 nucleotide sequences, respectively, and GenBank® Accession Nos. NP_001278413.1 and NP_082756.1 disclose exemplary human and mouse CEACAM5 protein sequences, respectively. One of ordinary skill in the art can identify additional CEACAM5 nucleic

5   acid and protein sequences, including CEACAM5 variants that retain CEACAM5 biological activity (such as having increased levels in urine from a subject with prostate cancer).

Carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6): Also known as non-specific cross-reacting antigen (NCA; *e.g.*, OMIM 163980), normal

10   cross-reacting antigen cea-like protein (CEAL), and cluster designation 66c (CD66c), CEACAM6 is a cell surface glycoprotein in the CEA family. CEACAM6 is expressed in neutrophils, affects tumor cell sensitivity to adenovirus infection, and is a receptor for *E. coli* adhesion to epithelial cells in patients with Crohn's disease. Further, CEACAM6 plays a role in platelet activation, signaling, and aggregation as well as cell surface

15   interactions at the walls of blood vessels. CEACAM6 has also been associated with colorectal and other cancers.

Includes CEACAM6 nucleic acid molecules and proteins. CEACAM6 sequences are publicly available. For example, GenBank® Accession Nos. NM_002483.6 and BC078962.1 disclose exemplary human and rat CEACAM6

20   nucleotide sequences, respectively, and GenBank® Accession Nos. NP_002474.4 and AAH78962.1 disclose exemplary human and rat CEACAM6 protein sequences, respectively. One of ordinary skill in the art can identify additional CEACAM6 nucleic acid and protein sequences, including CEACAM6 variants that retain CEACAM6 biological activity (such as having increased levels in urine from a subject with prostate

25   cancer).

Chemokine (C-C Motif) ligand 3 (CCL3): Also known as small inducible cytokine A3 (SCYA3; *e.g.,* OMIM 182283), macrophage inflammatory protein 1-α (MIP1α), and tonsillar lymphocyte LD78 α Protein (LD78-α), CCL3 is a monokine involved in the acute inflammatory state of polymorphonuclear leukocyte recruitment

30   and activation. CCL3 is expressed in many cell types, but most notably macrophages,

dendritic cells, and lymphocytes. Further, CCL3 plays a key role in inflammation and the immune response to infection and can promote homeostasis. CCL3 has been associated with such diseases as HIV and rheumatoid arthritis.

Includes CCL3 nucleic acid molecules and proteins. CCL3 sequences are

5   publicly available. For example, GenBank® Accession Nos. NM_002983.2, NM_013025.2, and NM_011337.2, disclose exemplary human, rat, and mouse CCL3 nucleotide sequences, respectively, and GenBank® Accession Nos. NP_002974.1, EDM05492.1, and NP_035467.1 disclose exemplary human, rat, and mouse CCL3 protein sequences, respectively. One of ordinary skill in the art can identify additional

10   CCL3 nucleic acid and protein sequences, including CCL3 variants that retain CCL3 biological activity (such as having increased levels in urine from a subject with prostate cancer).

Cysteine-rich secretory protein 3 (CRISP3): Also known as CRS3 and Aeg2, CRISP3 is in the cysteine-rich secretory protein subgroup of the CAP protein

15   superfamily, the subgroup members of which are substantially implicated in mammalian reproductive system function. CRISP3 is expressed in neutrophils, reproductive organs and glands, and the thymus and colon. CRISP3 is a glycoprotein that plays a role in endometrial remodeling and repair (*e.g.*, during the menstrual cycle and pregnancy), in prostate cancer, and immunity (*e.g.*, in hepatitis C and Sjögren's

20   syndrome).

Includes CRISP3 nucleic acid molecules and proteins. CRISP3 sequences are publicly available. For example, GenBank® Accession Nos. BC069602.1 and NM_009639.2 disclose exemplary human and mouse CRISP3 nucleotide sequences, respectively, and GenBank® Accession Nos. EAX04348.1 and AAI32539.1 disclose

25   exemplary human and mouse CRISP3 protein sequences, respectively. One of ordinary skill in the art can identify additional CRISP3 nucleic acid and protein sequences, including CRISP3 variants that retain CRISP3 biological activity (such as having increased levels in urine from a subject with prostate cancer).

CXC motif, ligand 14 (CXCL14): Also known as small inducible cytokine

30   subfamily B, member 14 (SCYB14; *e.g.,* OMIM 604186), CXC chemokine in breast

- 15 -

and kidney (BRAK), and MIP-2g, CXCL14 is a small cytokine in the CXC family.
CXCL14 is expressed at high levels in many normal tissues, but is notably absent in
many cancerous tissues. Further, CXCL14 plays a role in chemotaxis, homing, and
activation for cells involved in the immune response and has been shown to inhibit

5   angiogenesis.

Includes CXCL14 nucleic acid molecules and proteins. CXCL14 sequences are
publicly available. For example, GenBank® Accession Nos. NM_004887.4,
NM_001013137.2, and NM_019568.2 disclose exemplary human, rat, and mouse
CXCL14 nucleotide sequences, respectively, and GenBank® Accession Nos.

10   AAH03513.1, AAI01897.1, and AAH79661.1 disclose exemplary human, rat, and
mouse CXCL14 protein sequences, respectively. One of ordinary skill in the art can
identify additional CRISP3 nucleic acid and protein sequences, including CXCL14
variants that retain CXCL14 biological activity (such as having increased levels in urine
from a subject with prostate cancer).

15   **Differential expression or altered expression**: A difference, such as an
increase or decrease, in the conversion of the information encoded in a gene (such as a
prostate cancer-related gene) into messenger RNA, the conversion of mRNA to a
protein, or both. In some examples, the difference is relative to a control or reference
value, such as a cut-off value of expression for each marker. Detecting differential

20   expression can include measuring a change in gene or protein expression, such as a
change in expression of one or more prostate cancer-related genes or proteins disclosed
herein.

Interleukin 24 (IL24): Also known as suppression of tumorigenicity 16 (ST16;
e.g., OMIM 604136), melanoma differentiation-associated gene 7 (MDA7), ML-1, and

25   IL-17F, IL24 is a cytokine and tumor-suppressing protein in the IL-10 family of
cytokines. IL24 plays a role in cell survival and proliferation as well as wound healing,
psoriasis, and cancer. Further, LI24 is expressed by cells involved in the immune
response and can then act in skin, lung, and reproductive tissues.

Includes IL24 nucleic acid molecules and proteins. IL24 sequences are publicly

30   available. For example, GenBank® Accession Nos. NM_006850.3, NM_133311.1, and

NM_053095.2 disclose exemplary human, rat, and mouse IL24 nucleotide sequences, respectively, and GenBank® Accession Nos. NP_006841.1, NP_579845.1, and NP_444325.2 disclose exemplary human, rat, and mouse IL24 protein sequences, respectively. One of ordinary skill in the art can identify additional IL24 nucleic acid

5   and protein sequences, including IL24 variants that retain IL24 biological activity (such as having increased levels in urine from a subject with prostate cancer).

Isolated: An "isolated" biological component (such as a nucleic acid molecule, protein, or cell) has been substantially separated or purified away from other biological components in the cell of the organism, or the organism itself, in which the component

10   naturally occurs, such as other chromosomal and extra-chromosomal DNA and RNA, proteins and cells. Nucleic acid molecules and proteins that have been "isolated" include prostate cancer-related molecules (such as DNA or RNA) and proteins purified by standard purification methods. The term also embraces nucleic acid molecules, proteins and peptides prepared by recombinant expression in a host cell as well as

15   chemically synthesized nucleic acid molecules and proteins. For example, an isolated protein, such as a prostate cancer-related protein, is one that is substantially separated from other types of proteins in a cell.

Label: An agent capable of detection, for example by mass spectrometry, ELISA, spectrophotometry, flow cytometry, or microscopy. For example, a label can

20   be attached to a nucleic acid molecule or protein, thereby permitting detection of the nucleic acid molecule or protein. For example, a protein or peptide can be produced as a heavy, stable isotope, but as a protein or peptide with ^{13}C or ^{15}N incorporated as a heavy, stable isotope. Examples of labels include, but are not limited to, radioactive or heavy, stable isotopes, enzyme substrates, co-factors, ligands, chemiluminescent agents,

25   fluorophores, haptens, enzymes, and combinations thereof. Methods for labeling and guidance in the choice of labels appropriate for various purposes are discussed for example in Sambrook *et al*. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989) and Ausubel *et al*. (In Current Protocols in Molecular Biology, John Wiley & Sons, New York, 1998).

Matrix metalloproteinase 9 (MMP9): Also known as collagenase type IV B (CLG4B; *e.g.,* OMIM 120361), gelatinase B (GELB), collagenase type V, and 92-KD gelatinase, MMP9 is a 92-kD type IV collagenase and a member of the zinc metalloproteinase family. MMP9 is expressed in cells involved in the immune response

5   and has been found in skin, lung, and synovial tissues. Further, as a matrix metalloproteinase, MMP9 aids in breaking down the extracellular matrix for normal physiological processes; MMP9 is involved in the immune response, angiogenesis, and wound repair and is associated with autoimmune diseases, cancer, and cardiovascular disease.

10   Includes MMP9 nucleic acid molecules and proteins. MMP9 sequences are publicly available. For example, GenBank® Accession Nos. NM_004994.2, NM_031055.1, and NM_013599.4 disclose exemplary human, rat, and mouse MMP9 nucleotide sequences, respectively, and GenBank® Accession Nos. EAW75776.1, EDL96479.1, and EDL06438.1 disclose exemplary human, rat, and mouse MMP9

15   protein sequences, respectively. One of ordinary skill in the art can identify additional MMP9 nucleic acid and protein sequences, including MMP9 variants that retain MMP9 biological activity (such as being increased in the urine of prostate cancer patients, particularly intermediate- to high-risk prostate cancer).

Periostin (POSTN): Also known as PN (*e.g.,* OMIM 608777) and osteoblast-

20   specific factor 2 (OSF2), POSTN is an extracellular matrix protein. POSTN is expressed in many normal tissues, including aortis, stomach, lower gastrointestinal tract, placental, uterine, and breast tissues. Further, POSTN plays a role in tissue development and regeneration as well as epithelial cell adhesion and migration; POSTN is involved in cancer stem cell maintenance and metastasis.

25   Includes POSTN nucleic acid molecules and proteins. POSTN sequences are publicly available. Nucleic acid and protein sequences for POSTN are publicly available. For example, GenBank® Accession Nos. BC106709.1, NM_001108550.1, and BC031449.1 disclose exemplary human, rat, and mouse POSTN nucleotide sequences, respectively, and GenBank® Accession Nos. AAI06710.1,

30   NP_001102020.1, and AAH31449.1 disclose exemplary human, rat, and mouse POSTN

protein sequences, respectively. One of ordinary skill in the art can identify additional POSTN nucleic acid and protein sequences, including POSTN variants that retain POSTN biological activity (such as being increased in the urine of prostate cancer patients).

5   **Prostate-specific antigen (PSA)**: Also known as kallikrein-related peptidase 3 (kallikrein 3, KLK3; e.g., OMIM 176820); antigen, prostate-specific (APS); and gamma-seminoprotein, PSA is a glycoprotein and member of the kallikrein-related peptidase family. PSA is predominantly secreted by epithelial cells in the prostate gland and functions to dissolve cervical mucus to facilitate sperm entry into the uterus. PSA is
10   a controversial tool used to diagnose prostate cancer, as increased PSA levels in blood may suggest the presence of prostate cancer.

         Includes PSA nucleic acid molecules and proteins. PSA sequences are publicly available. Nucleic acid and protein sequences for PSA are publicly available. For example, GenBank® Accession Nos. NM_001648.2, NM_012725.2, and
15   NM_008455.3 discloses exemplary human PSA nucleotide sequences, respectively, and GenBank® Accession Nos. CAD54617.1, AAH89815.1, and NP_001639.1 discloses exemplary human PSA protein sequences. One of ordinary skill in the art can identify additional PSA nucleic acid and protein sequences, including PSA variants that retain PSA biological activity (such as being secreted by the prostate gland).

20   **Prostate cancer**: Also known as carcinoma of the prostate, prostate cancer is the development of cancer in the prostate, a gland in the male reproductive system. Prostate cancer is classified into different risk categories, including low-, intermediate-, and high-risk prostate cancer, which means that a patient has a low-, intermediate-, and high-risk, respectively, of pathological and biochemical outcomes after radical
25   prostatectomy; metastasis; prostate cancer-specific mortality; and all-cause mortality (Cooperberg *et al.*, J Cancer Inst., 101(12):878-887, 2009). One means of assessing the risk is using Gleason scoring: low-risk prostate cancer, Gleason score sum less than or equal to 6; intermediate-risk prostate cancer, Gleason score sum at 7; and high-risk prostate cancer, Gleason score sum greater than 7. Most prostate cancers are slow
30   growing; however, some grow relatively quickly. The cancer cells may spread from the

prostate to other parts of the body, particularly the bones and lymph nodes. It may initially cause no symptoms. In later stages, it can lead to difficulty urinating, blood in the urine, or pain in the pelvis, back or when urinating or to feeling tired due to low levels of red blood cells.

5   Prostate cancer can be diagnosed by biopsy. Medical imaging may then be done to determine if the cancer has spread to other parts of the body. Prostate cancer screening is controversial. Prostate-specific antigen (PSA) testing increases cancer detection but does not decrease mortality. The United States Preventive Services Task Force recommends against screening using the PSA test, due to the risk of overdiagnosis

10   and overtreatment, as most cancer diagnosed would remain asymptomatic, and concludes that the potential benefits of testing do not outweigh the expected harms.

Many cases can be safely followed with active surveillance or watchful waiting. Other treatments may include a combination of surgery (such as cryotherapy), radiation therapy, hormone therapy, and chemotherapy. When it only occurs inside the prostate it

15   may be curable. In those in whom the disease has spread to the bones, pain medications, bisphosphonates and targeted therapy, among others, may be useful. Outcomes depend on a person's age and other health problems as well as how aggressive and extensive the cancer is. Most people with prostate cancer do not die from the disease. The 5-year survival rate in the United States is 99%. Globally, it is the second

20   most common type of cancer and the fifth leading cause of cancer-related death in men. Studies of males who died from unrelated causes have found prostate cancer in 30% to 70% of those over age 60.

Sample: A biological specimen containing genomic DNA, RNA (*e.g.*, mRNA), protein, or combinations thereof, obtained from a subject. Examples include, but are not

25   limited to, peripheral blood, serum, plasma, urine, saliva, tissue biopsy, fine needle aspirate, surgical specimen, and autopsy material. In one example, a sample is a urine sample from a subject with or at risk for prostate cancer, such as low-, intermediate-, or high-risk prostate cancer. In some examples, samples are used directly in the methods provided herein. In some examples, samples are manipulated prior to analysis using the

30   disclosed methods, such as through concentrating, filtering, centrifuging, diluting,

desalting, denaturing, reducing, alkylating, proteolyzing, or combinations thereof. In some examples, components of the samples are isolated or purified prior to analysis using the disclosed methods, such as isolating cells, proteins, and/or nucleic acid molecules from the samples.

5   **Secreted frizzled-related protein 4 (SFRP4)**: Also known as frizzled-related protein (*e.g.*, OMIM 606570) and human endometrium (FRPHE), SFRP4 is in the SFRP family, the members of which regulate Wnt signaling. SFRP4 is expressed in the endometrium, myocardium, breast tissue, and islets. Further, SFRP4 plays a role in regulating apoptosis, insulin secretion, and in regulating uterine morphology and

10   function. SFRP4 is associated with bone diseases, such as rickets and bone cancers.

Includes SFRP4 nucleic acid molecules and proteins. SFRP4 sequences are publicly available. For example, GenBank® Accession Nos. NM_003014.3, NM_053544.1, and NM_016687.3 disclose exemplary human, rat, and mouse SFRP4 nucleotide sequences, respectively, and GenBank® Accession Nos. CAG46532.1,

15   NP_445996.1, and AAH34853.1 disclose exemplary human, rat, and mouse SFRP4 protein sequences, respectively. One of ordinary skill in the art can identify additional SFRP4 nucleic acid and protein sequences, including SFRP4 variants that retain SFRP4 biological activity (such as being increased in the urine of prostate cancer patients).

Subject: Living multi-cellular vertebrate organisms, a category that includes

20   mammals, such as human and non-human mammals, such as veterinary subjects (for example cats, dogs, cows, sheep, horses, pigs, and mice). In a particular example, a subject is one who has or is at risk for prostate cancer, such as low-, intermediate-, or high-risk prostate cancer. In a particular example, a subject is one who is suspected of having prostate cancer.

25   **Therapeutically effective amount**: An amount of a pharmaceutical preparation that alone, or together with a pharmaceutically acceptable carrier or one or more additional therapeutic agents, induces the desired response. A therapeutic agent, such as an anti-neoplastic chemotherapeutic agent, radiotherapeutic agent, or biologic agent, is administered in therapeutically effective amounts.

Therapeutic agents can be administered in a single dose, or in several doses, for example daily, during a course of treatment. However, the effective amount can be dependent on the source applied, the subject being treated, the severity and type of the condition being treated, and the manner of administration. Effective amounts of a

5   therapeutic agent can be determined in many different ways, such as assaying for a sign or a symptom of an adenocarcinoma. Effective amounts also can be determined through various *in vitro*, *in vivo* or *in situ* assays. For example, a pharmaceutical preparation can decrease one or more symptoms of a prostate cancer, for example, a decrease in the size of the prostate cancer, the number of tumors, the number of

10   metastases, or other symptoms (or combinations thereof) by at least 20%, at least 50%, at least 70%, at least 90%, at least 98%, or even 100%, as compared to an amount in the absence of the pharmaceutical preparation.

Treating a disease: "Treatment" refers to a therapeutic intervention that ameliorates a sign or symptom of a disease or pathological condition, such a sign or

15   symptom of prostate cancer. Treatment can also induce remission or cure of a condition, or can reduce the pathological condition, such as a reduction in tumor size, a reduction in tumor burden, a reduction in a sign or a symptom of a tumor (such as cachexia), a reduction in metastasis, or combinations thereof. In particular examples, treatment includes preventing a disease, for example by inhibiting the full development

20   of a disease, such as decreasing the ability of a tumor to metastasize. Prevention of a disease does not require a total absence of disease.

Upregulated or activation: When used in reference to the expression of a nucleic acid molecule, such as a gene, refers to any process which results in an increase in the production of a gene product. A gene product can be RNA (such as mRNA,

25   rRNA, tRNA, and structural RNA) or protein. Therefore, gene upregulation or activation includes processes that increase transcription of a gene or translation of mRNA, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4.

Examples of processes that increase transcription include those that facilitate

30   formation of a transcription initiation complex, those that increase transcription

- 22 -

initiation rate, those that increase transcription elongation rate, those that increase

processivity of transcription, and those that relieve transcriptional repression (for

example, by blocking the binding of a transcriptional repressor). Gene upregulation can

include inhibition of repression as well as stimulation of expression above an existing

5   level. Examples of processes that increase translation include those that increase

translational initiation, those that increase translational elongation and those that

increase mRNA stability.

Gene upregulation includes any detectable increase in the production of a gene

product, such as a protein. In certain examples, production of a gene product increases

10   by at least 2-fold, at least 3-fold, at least 4-fold, or at least 5-fold, as compared to a

control.

Overview

Provided herein are 14 prostate cancer-associated secreted proteins that are

15   found in increased levels in the urine of patients with prostate cancer, as compared to

levels in urine from patients without prostate cancer. Sensitive multiplexed assays were

developed for reliable simultaneous quantification of 12 detectable secreted proteins in

human urine. Except for CCL3 and POSTN, the other 10 proteins were reproducibly

detected and quantified in all the urine samples analyzed with at least one surrogate

20   peptide. The peptide signatures from AGR2, AGR3, CEACAM5, CD90 and SFRP4, in

particular, produced an area-under-the curve (AUC) of 0.95. It was also observed that

urinary MMP9 levels increased with higher risk cancer, which correlated with increase

in MMP9 gene expression in Gleason pattern ≥ 7 *vs*. Gleason ≤ 6 prostate cancer subjects

Samples of voided urine (≤ 90 ml) were collected without DRE from non-cancer

25   controls, pre-op(erative), and post-op patients; post-pellet urine supernatant was spin

concentrated (to <1 ml); urinary proteins were simultaneously measured by sensitive

multiplex targeted proteomics. The quantification accuracy of each surrogate peptide

assay was evaluated by correlation analysis of relative protein abundance between

surrogate peptides from the same protein. The ability of measured markers to

30   distinguish cancer from non-cancer was assessed by AUC values.

To date, disease detection relies mostly on single markers. With each additional marker, more information is obtained, which provides a corresponding increase in the area under the curve (AUC), which is the probability that the biomarker will rank a randomly chosen positive diagnosis higher than a randomly chosen negative one (with a

5   perfect biomarker, the AUC is 1.0, but a biomarker with no practical utility produces an AUC of 0.5). The secreted protein cancer-related biomarkers disclosed herein can generate an AUC over 0.9 for detecting prostate cancer. These markers were identified from their elevated expression in cell types of prostate tumors. Multimarker measurement can be accurately obtained, for example by using proteomics tools. Assay

10   development was rapid for finding the right surrogate peptides and implementation compared to immunoassays that require the time consuming generation of high quality monoclonal antibodies and their testing. With continuous advancements in measurement sensitivity, such as through LG-SRM [13] or PRISM-SRM [14], SRM assays are useful for measuring protein biomarkers in tissues [32,33] and human body

15   fluids [9,34] as well as for facilitating the transition of biomarkers to large-scale clinical validation trials.

Accuracy in SRM assays depends on the surrogate peptides selected for each biomarker; the measured concentrations are directly proportional to the cognate protein analytes. As shown by the PSA surrogate peptides examined herein, unknown

20   modifications on the analyte peptide sequences produce lower levels of measured concentrations compared with the true concentrations. Without modifications, the abundance ratio for any two or more surrogate peptides is constant across multiple samples analyzed, and a good correlation curve with no significant data point deviation is produced. Based on the correlation coefficients, the quantification accuracy of

25   individual selected surrogate peptides of the secreted protein markers was assured, and the surrogate peptides were selected for SRM assay configuration for all subsequent testing. In studies involving human cell lines, most surrogate peptides (453/466) generated a high correlation coefficient (R^2 >0.8) with no significant data point deviation [25]. Many surrogate peptides used in the urine analysis generated lower

30   correlation coefficients (the median $R^2 = 0.70$) with data point deviations. This

- 24 -

observation shows that the targeted protein molecules in urine of multiple patient samples are more varied, such as due to allelic differences or cancer-associated isoforms, than those in single cell lines. PSA measurement in urine, for example, was shown to be dependent on the peptide chosen. A single PSA peptide could not

5   accurately quantify PSA in three of the urine samples (FIG. 8, urine IDs. P06017Pre, P07040Pre, and P08015Pre). A specific PSA proteoform encoded by SNP-L132I (rs2003783) is located within the LSEPAE(L/I)TDAVK surrogate peptide. This newly identified PSA proteoform was observed in 9 out of 72 clinical serum samples [27], and in the P06-017C urine. Abundance recalculations based on this information improved

10   the correlation for this urine sample.

In addition to surrogate peptide selection, PSA normalization was used to account for the contribution of protein markers from sources other than the prostate. Without PSA normalization, the median values of the marker protein L/H ratios were lower in the non-cancer urine than the cancer urine, while, for CRISP3 and SFRP4, the

15   median values in the non-cancer urine were higher than the cancer urine. For example, the L/H ratio for WANQCNYR/CRISP3 in cancer was 0.29, and 0.38 in non-cancer. Given that CRISP3 and SFRP4 are produced by other sources in the urinary tract, this confounding result was not unexpected. Expression of CRISP3 and the non-secreted prostate cancer marker AMACR is localized to other tissues in the urinary tract [35,36].

20   A similar finding for AGR2 in urine was obtained [9,30], even though a database query showed no detectable AGR2 in normal urine. One set of urine samples produced a 0.002 L/H ratio in cancer compared to 0.005 in non-cancer [9]. AGR2 can originate from several other sources, such as the bladder and kidney (moderate AGR2 immunostaining of tubular cells, unpublished data). Although normal AGR2-expressing

25   bladder urothelial cells do not secrete AGR2 [30], unknown physiological factors could produce detectable levels of AGR2 in urine. While urinary CRISP3 and SFRP4 are constitutively produced by non-prostate sources, urinary AGR2 is transiently produced from non-prostate sources. Urinary PSA, on the other hand, is exclusively produced by the prostate. The significantly higher concentrations of urinary PSA found in some non-

30   cancer samples could be due to donors with an enlarged prostate from benign

hyperplasia. For example, prostate cancer patients with a prostate volume of 35 cm^3 (n = 29) and benign prostatic hyperplasia patients prostate volume of 45 cm^3 (n = 35) were measured to have median urinary PSA levels of 52.6 ng/mL and 123.2 ng/mL, respectively [37]. With PSA normalization, the performance of the disclosed prostate

5   cancer protein biomarkers showed a near perfect AUC.

The use of the disclosed prostate cancer-related biomarkers reduces the need for prostate biopsy, an invasive, expensive, and potentially morbid procedure with up to a 4% risk of sepsis [38]. In one example, prostate cancer diagnosis includes the use of a the disclosed prostate cancer-related markers before or following an abnormal serum

10   PSA result. Urine donation is convenient and does not require a DRE. With regard to the possibility that DRE might enhance marker signals, the levels of two post-DRE urine, P08-032C and P08-036N, were obtained. The signals from P08-032C (Gleason 4+3, tumor volume 1 cc) were not higher than those from others obtained without a DRE. DRE did not produce any increased signals in non-cancer P08-036N. With

15   regard to possible age-related increases in marker levels, samples from a 76-year-old man (P08-018N) and a 53-year-old man (P08-022N) were analyzed. There was no detected increase in the background urinary levels of these proteins. It is possible that the baseline level of these markers may remain more or less constant with age. In contrast, an increase in serum PSA with age is known [39]. In one example, if the

20   prostate cancer-related protein panel result is negative (*e.g.*, the analyzed proteins are not increased in the urine), no biopsy would be necessary if the negative predictive value is sufficiently high. Furthermore, the ratio of MMP9/PSA concentrations can be used to distinguish low volume/low grade prostate cancer from significant cancer. Therefore, effectively integrating the results from analyzing expression of the disclosed

25   prostate cancer-related molecules will lead to greater detection of significant cancer with fewer biopsies performed in patients without cancer.

For AGR2 quantification, both mass spectrometry proteomics and ELISAs were used with good correlation, $R^2 = 0.91$ [9], and similar AUC values (0.75). Thus, the disclosed prostate cancer-related biomarkers can be used for a multiplex ELISA to

30   measure all these proteins simultaneously in a clinical setting. The equivalency in

urinary AGR2 quantification by PRISM-SRM and ELISA shows that a multiplex ELISA can replicate the proteomics results, and clinical proteomics can make large-scale testing feasible.

5   **Evaluating Expression in a Subject with a Risk of Prostate Cancer**

Provided herein are methods of diagnosing a subject with a risk of prostate cancer and methods of treating a subject with prostate cancer (such as a human or veterinary subject). In particular examples, the methods can determine with high specificity, sensitivity, and accuracy (such as having an AUC of greater than 0.8,

10   including, for example, at least 0.85, at least 0.9, or at least 0.95) whether a subject is likely to have prostate cancer. The prostate cancer can be any risk category of interest, including low- (Gleason score sum is 6 or lower), intermediate- (Gleason score sum is 7), and high-risk (Gleason score sum is above 7) prostate cancer. It is helpful to be able to determine whether or not a subject has prostate cancer because there are a variety of

15   protocols for diagnosing prostate cancer but not all are specific, sensitive, and accurate. Hence, using the results of the disclosed assays to help distinguish subjects that are likely to have prostate cancer versus those not likely to have prostate cancer offers a substantial clinical benefit and allows subjects to be accurately diagnosed and, if a subject has prostate cancer, to be accurately treated.

20   In additional examples, the methods are utilized to determine whether or not to provide the subject with therapeutic intervention. In one example, a therapeutic intervention is administered. Thus, if the subject has prostate cancer, a therapeutic intervention, such as watchful waiting, active surveillance, surgery, radiation, hormone therapy, chemotherapy, brachytherapy, cryotherapy, ultrasound, bisphosphate therapy,

25   biologic therapy, or vaccine therapy can be utilized. Using the results of the disclosed assays to help distinguish subjects that are likely to have prostate cancer versus those not likely to have prostate cancer offers a substantial clinical benefit because, where the subject has prostate cancer, the methods disclosed herein allow the subject to be selected for therapeutic intervention.

Methods of diagnosing a subject with a risk of prostate cancer and methods of treating a subject with prostate cancer, such as low-, intermediate-, or high-risk prostate cancer, are provided. Such methods can include measuring or detecting absolute or relative amounts of prostate cancer-related markers present in a sample (such as a urine

5   sample) obtained from the subject, for example, using surrogate peptides of the marker proteins (*e.g.*, as shown in FIGS. 4, 12 and 15) and/or antibodies, nucleic acid probes, and/or nucleic acid primers specific for each marker. In some examples, the prostate cancer-related markers can include at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11 or all 12 of (such as 3, 4, 5, 6, 7, 8, 9, 10, 11 or

10   12 of) AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4. The expression levels of these markers can be measured. If increased protein and/or nucleic acid expression of the prostate cancer-related markers, for example, expression of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, in the

15   sample is measured, the method can include administering therapeutic intervention to the subject, thereby treating the subject.

In some examples, measuring expression of prostate cancer-related markers, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, can include quantitating protein and/or nucleic

20   acid expression of these markers in a sample obtained from the subject. In particular examples, these markers are first analyzed for measurement accuracy, such as correlating the amounts of different surrogate peptides from the same prostate cancer-related marker protein where the protein expression is measured.

In other examples, measuring increased protein or nucleic acid expression of the

25   markers AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 is relative to an amount of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 median protein or nucleic acid expression, respectively, for example a median value of protein or nucleic acid expression for each marker expected in a subject

30   with no prostate cancer.

In some examples, measuring protein and/or nucleic acid expression of the markers AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 can include measuring more than one marker, such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or all 12 of the markers. In other examples, any combination of these markers can be measured. In particular examples, any of the combinations of markers listed in FIG. 15 can be measured.

In some examples, measuring expression of prostate cancer-related markers, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, can include measuring the amount of protein expressed. For example, measuring the amount of protein expressed can include measuring a surrogate peptide from the protein. In some specific examples, surrogate peptides can be generated through contacting the protein with a protease, such a trypsin. In some particular examples, surrogate peptides for the prostate cancer-related markers AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 include the surrogate peptides listed in FIG. 4. In other particular examples, combinations of surrogate peptides for prostate cancer-related markers can be used, such as the combinations listed in FIG. 15.

In some examples, measuring the amount of protein expressed for prostate cancer-related markers, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, can include using mass spectroscopy and/or an immunoassay.

In particular examples, measuring the amount of protein expressed can include measuring the protein concentration using an immunoassay. In particular examples, the immunoassay can be an ELISA.

In some examples, measuring expression of prostate cancer-related markers, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, using the amount of protein expressed can include measuring the protein concentration using mass spectrometry. In some examples, mass spectrometry can be used to determine the protein concentration of the full-length protein and/or surrogate peptide(s) for the protein. In particular examples,

mass spectrometry can be used to determine the protein concentration of prostate cancer-related markers, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, using surrogate peptides, such as the peptides listed in FIG. 4.

5   In particular examples, measuring expression of prostate cancer-related markers by using mass spectrometry can include using mass spectrometry assays such as LC-SRM, LG-SRM, and/or PRISM-SRM. In some examples, measuring expression of prostate cancer-related markers (such as in a serum sample) can include using an LC-SRM assay, for example, where the serum protein levels are least at a moderate

10   abundance, such as about low µg/mL (*e.g.*, 1-10, 10-50, 50-100, or 100-500 µg/mL).

In other examples, the measuring increased protein or nucleic acid expression of the markers AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 includes measuring some markers that are at low-to-moderate abundance, for example, in the range of about low µg/mL to high

15   ng/mL (*e.g.*, 1-10 µg/mL, 500 ng/mL-1 µg/mL, or 100-500 ng/mL), in the sample obtained from the subject. In particular examples, the low-to-moderate-abundance markers in the sample can include CRISP3, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4. In some examples, these low-abundance markers can be accurately measured by using assays with sufficient sensitivity, such as an LG-SRM assay, a

20   PRISM-SRM assay, and/or an ELISA.

In certain examples, measuring increased protein or nucleic acid expression of the markers AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 includes measuring some markers that are at low abundance, for example, in the range of about low ng/mL to high pg/mL (*e.g.*, 500-

25   100 ng/mL, 100-50 ng/mL, 50-10 ng/mL, 10-1 ng/mL, 500 pg/mL-1 ng/mL, 500-100 pg/mL, or 100-50 pg/mL) in the sample obtained from the subject. In particular examples, the low-abundance markers in the sample can include AGR2, AGR3, CCL3, CEACAM5, and CEACAM6. In some examples, these low-abundance markers can be accurately measured by using assays with sufficient sensitivity, such as a PRISM-SRM

30   assay and/or an ELISA.

- 30 -

In other examples, measuring increased protein or nucleic acid expression of prostate cancer-related markers, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, includes normalizing expression of the prostate cancer-related molecules to the expression of a prostate molecule. In particular examples, the prostate molecule used for normalization can be PSA. In specific examples, normalizing to the PSA concentration can include normalizing the protein expression of a prostate cancer-related marker to the amount of PSA protein. In other examples, normalizing to the amount of PSA can include normalizing the protein expression of a prostate cancer-related marker to the amount of at least one surrogate peptide from the PSA protein. In certain examples, the PSA protein surrogate peptide(s) can include IVGGWECEK, LSEPAELTDAVK, or both.

In some examples, the methods can include measuring increased expression of two or more prostate cancer-related markers, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, which can include the combinations of markers listed in FIG. 15. In particular examples, the protein expression of the markers can be measured, for example, by using surrogate peptides of the markers, such as the surrogate peptides listed in FIG. 4. In specific examples, the amounts of surrogate peptides can be measured using mass spectrometry. In some examples, the mass spectrometry technique used to measure the amounts of surrogate peptides can be LC-SRM, LG-SRM, and/or PRISM-SRM. In another example, the amounts of surrogate peptides are normalized to a prostate molecule, such as PSA, for example, by using one or more surrogate peptides of PSA (*e.g.,* IVGGWECEK, LSEPAELTDAVK, or both). In specific examples, the increased expression measured for the at least two prostate cancer-related markers (such as 3, 4 5, 6, 7, 8, 9, 10, 11 or all 12 of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4) has an AUC value greater than 0.80, such as an AUC value of 0.81, 0.82, 0.83, 0.84, 0.85, 0.86, 0.87, 0.88, or 0.89. In other specific examples, the increased expression measured for the at least two prostate cancer-related markers (such as 3, 4 5, 6, 7, 8, 9, 10, 11 or all 12 of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and

SFRP4) has an AUC value greater than 0.90, such as an AUC value of 0.91, 0.92, 0.93, 0.94, 0.95, 0.96, 0.97, 0.98, or 0.99.

      Methods of diagnosing intermediate- or high-risk prostate cancer and treating a subject with intermediate- or high-risk prostate cancer are provided. In particular

5   examples, the methods can determine with significant accuracy whether a subject is likely to have prostate cancer associated with a specific risk category, such as low-, intermediate-, and high-risk prostate cancer. In particular examples, the methods can distinguish with significant accuracy between subjects that have low-risk prostate cancer and subjects that have intermediate-to-high-risk prostate cancer. It is helpful to be able

10   to determine whether or not a subject has intermediate-to-high-risk prostate cancer because there are a variety of protocols for treating prostate cancer. Hence, using the results of the disclosed assays to help distinguish subjects that are likely to have prostate cancer associated with a specific risk category offers a substantial clinical benefit and allows subjects to be accurately diagnosed and, if a subject has prostate cancer

15   associated with a specific risk category, to be accurately treated.

      In some examples, methods of diagnosing intermediate- or high-risk prostate cancer and treating a subject with intermediate- or high-risk prostate cancer can include measuring a sample (such as a urine sample) obtained from the subject for example, using surrogate peptides of the marker proteins and/or antibodies, nucleic acid probes,

20   and/or nucleic acid primers specific for the markers. In some examples, the prostate cancer-related markers can include MMP9. The expression levels of the markers can be measured. If increased protein and/or nucleic acid expression of one or more prostate cancer-related markers, for example, expression of MMP9, in the sample is measured, the method can include administering therapeutic intervention for intermediate- or high-

25   risk prostate cancer to the subject, thereby treating the subject.

      In particular examples, expression of MMP9 in a sample obtained from a subject is measured. In one example, the expression of MMP9 can be determined based on MMP9 protein concentration. In some examples, the expression of MMP9 protein can be measured using at least one surrogate peptide of MMP9. In specific examples, the

30   surrogate peptide can be FQTFEGDLK, LGLGADVAQVTGALR, or both.

In further examples, measuring increased expression of MMP9 in the sample obtained from the subject can include comparing the expression of MMP9, such as by using the concentration of the MMP9 surrogate peptide(s) FQTFEGDLK and/or LGLGADVAQVTGALR, to the amounts of MMP9 expression expected in a sample

5   from a subject who has low-risk prostate cancer. In some specific examples, the surrogate peptides FQTFEGDLK and LGLGADVAQVTGALR for MMP9 protein can be used to measure expression of MMP9 with significant accuracy. In certain examples, the expression of MMP9 is increased compared to a sample from a subject who has low-risk prostate cancer, such as where the expression of MMP9 is increased compared with

10   a sample from a subject who has low-risk prostate cancer with a P value of less than 0.05, such as a P value of 0.022 or a range of P values between 0.01-0.02, 0.02-0.03, 0.03-0.04, or 0.05.

In some examples, where an increase in the expression of MMP9 is measured in the sample obtained from the subject compared with the expression of MMP9 expected

15   in a sample from a subject who has low-risk prostate cancer, the methods include administering treatment for intermediate- or high-risk prostate cancer, thereby treating the subject. For example, a patient with low-risk prostate cancer may not exhibit increased expression of MMP9 that exceeds the expression expected from a patient with low-risk prostate cancer and, therefore, may not necessarily be a good candidate for

20   invasive treatments and/or treatments with potentially harmful side effects but, rather, may be a good candidate for watchful waiting or active surveillance. In another example, a patient with intermediate-to-high-risk prostate cancer may exhibit increased levels of MMP9 compared with the expression expected from a patient with low-risk prostate cancer and, therefore, may be better candidate for treatments such as surgery,

25   radiation therapy, and hormone therapy but may not be a good candidate for observation-based treatments, such as watchful waiting or active surveillance.

Evaluating Nucleic Acid Expression

In some examples, expression of nucleic acids (*e.g.*, mRNA) of prostate cancer-

30   related markers, such as the markers AGR2, AGR3, CRISP3, CCL3, CEACAM5,

CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, are analyzed and, in some examples, quantified. Suitable biological samples can include urine, blood, plasma, or serum samples obtained from a subject having or a subject at risk for prostate cancer (such as intermediate- or high-risk prostate cancer). An increase in the amount

5   of nucleic acid molecules for the prostate cancer-related markers, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, in the sample indicates that the subject has prostate cancer and/or has intermediate- or high-risk prostate cancer, as described herein. In some examples, expression of the prostate cancer-related nucleic acid molecule is normalized

10   to PSA expression in the sample (such as by measuring PSA cDNA, genomic DNA, or mRNA in the urine sample). In some examples, the assay is multiplexed, in that expression of several nucleic acids are detected simultaneously or contemporaneously (Quek *et al*., *Prostate* 75:1886-95, 2015).

RNA can be isolated from a sample from a subject having or a subject at risk for

15   prostate cancer or for intermediate- or high-risk prostate cancer, such as a urine, blood, plasma, or serum sample, using methods well-known to one skilled in the art, including commercially available kits. General methods for mRNA extraction are well-known in the art and are disclosed in standard textbooks of molecular biology, including Ausubel *et al*., Current Protocols of Molecular Biology, John Wiley and Sons (1997). In one

20   example, RNA isolation can be performed using a purification kit, buffer set, and protease from commercial manufacturers, such as QIAGEN®, according to the manufacturer's instructions. For example, total RNA from cells in culture (such as those obtained from a subject) can be isolated using QIAGIN® RNeasy mini-columns. Other commercially available RNA isolation kits include MASTERPURE® Complete

25   DNA and RNA Purification Kit (EPICENTRE® Madison, Wis.) and Paraffin Block RNA Isolation Kit (Ambion, Inc.). Total RNA from tissue samples can be isolated using RNA Stat-60 (Tel-Test). RNA prepared from a biological sample can be isolated, for example, by cesium chloride density gradient centrifugation.

Methods of gene expression profiling include methods based on hybridization

30   analysis of polynucleotides, methods based on sequencing of polynucleotides, and other

methods in the art. In some examples, mRNA expression in a sample is quantified using northern blotting or in situ hybridization (Parker & Barnes, Methods in Molecular Biology 106:247-283, 1999); RNAse protection assays (Hod, *Biotechniques* 13:852-4, 1992); and PCR-based methods, such as reverse transcription polymerase chain reaction

5   (RT-PCR) (Weis *et al.*, *Trends in Genetics* 8:263-4, 1992). Alternatively, antibodies can be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. Representative methods for sequencing-based gene expression analysis include Serial Analysis of Gene Expression (SAGE) and gene expression analysis by massively parallel signature

10   sequencing (MPSS).

Methods for quantitating mRNA are well-known in the art. In one example, the method utilizes RT-PCR. For example, extracted RNA can be reverse-transcribed and the derived cDNA used as a template in the subsequent PCR reaction. A variation of RT-PCR that can be used is real time quantitative RT-PCR, which measures PCR

15   product accumulation through a dual-labeled fluorigenic probe (e.g. TAQMAN® probe).

Although the PCR step can use a variety of thermostable DNA-dependent DNA polymerases, it typically employs the Taq DNA polymerase. TaqMan® PCR typically utilizes the 5'-nuclease activity of Taq or Tth polymerase to hydrolyze a hybridization

20   probe bound to its target amplicon, but any enzyme with equivalent 5' nuclease activity can be used. Two oligonucleotide primers are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide, or probe, is designed to detect nucleotide sequence located between the two PCR primers. The probe is non-extendible by Taq DNA polymerase enzyme and is labeled with a reporter fluorescent dye and a quencher

25   fluorescent dye. Any laser-induced emission from the reporter dye is quenched by the quenching dye when the two dyes are located close together as they are on the probe. During the amplification reaction, the Taq DNA polymerase enzyme cleaves the probe in a template-dependent manner. The resultant probe fragments disassociate in solution, and the signal from the released reporter dye is free from the quenching effect

30   of the second fluorophore. One molecule of reporter dye is liberated for each new

molecule synthesized, and detection of the unquenched reporter dye provides the basis for quantitative interpretation of the data.

In some examples, gene expression is identified or confirmed using the microarray technique. In this method, nucleic acid sequences of interest (such as mRNA, cDNAs and oligonucleotides) are plated or arrayed on a microchip substrate. The arrayed sequences are then hybridized with specific DNA probes from biological sample of interest. In a specific embodiment of the microarray technique, PCR amplified inserts of cDNA clones are applied to a substrate in a dense array. Probes for the nucleotide sequences of prostate cancer-related markers, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, are applied to the substrate, and the array can consist essentially of or consist of these sequences or of certain combinations of sequences for these markers, such as the marker combinations listed in FIG. 15. The microarrayed nucleic acids are suitable for hybridization under stringent conditions. Fluorescently labeled cDNA probes may be generated through incorporation of fluorescent nucleotides by reverse transcription of RNA extracted from tissues of interest. Labeled cDNA probes applied to the chip hybridize with specificity to each spot of DNA on the array. After stringent washing to remove non-specifically bound probes, the chip is analyzed. Quantitation of hybridization of each arrayed element allows for assessment of corresponding mRNA abundance. With dual color fluorescence, separately labeled cDNA probes generated from two sources of RNA are hybridized pairwise to the array. The relative abundance of the transcripts from the two sources corresponding to each specified gene is thus determined simultaneously. Such methods have the sensitivity required to detect rare transcripts, which are expressed at a few copies per cell, and to reproducibly detect at least approximately two-fold differences in the expression levels (Schena *et al.*, *Proc. Natl. Acad. Sci. USA* 93(2):10614-9, 1996). Microarray analysis can be performed by commercially available equipment, following manufacturer's protocols, such as are supplied with Affymetrix GenChip technology or Incyte's microarray technology.

Serial analysis of gene expression (SAGE) is another method that allows the simultaneous and quantitative analysis of a large number of gene transcripts without the

- 36 -

need for providing an individual hybridization probe for each transcript. First, a short sequence tag (about 10-14 base pairs) is generated that contains sufficient information to uniquely identify a transcript, provided that the tag is obtained from a unique position within each transcript. Then, many transcripts are linked together to form long serial

5   molecules that can be sequenced, revealing the identity of the multiple tags simultaneously. The expression pattern of any population of transcripts can be quantitatively evaluated by determining the abundance of individual tags and identifying the gene corresponding to each tag. For more details see, for example, Velculescu *et al*., *Science* 270:484-7, 1995, and Velculescu *et al*., *Cell* 88:243-51, 1997.

10

Evaluating Protein Expression

In some examples, protein expression of prostate cancer-related markers, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, is analyzed and, in some examples, quantified. Suitable

15   biological samples include urine, blood, plasma, and/or serum samples obtained from a subject having or a subject at risk for prostate cancer, such as for intermediate- or high-risk prostate cancer. An increase in the amount of prostate cancer-related marker proteins, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins, in the sample indicates that

20   the subject has prostate cancer and/or has intermediate- or high-risk prostate cancer, as described herein. In some examples, expression of the prostate cancer-related protein is normalized to PSA expression in the sample (such as by measuring a surrogate protein(s) for PSA in the urine sample). In some examples, the assay is multiplexed, in that expression of several proteins is detected simultaneously or contemporaneously.

25   The expression of prostate cancer-related markers, such as 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, can be measured using any of a number of techniques, such as direct physical measurements (*e.g*., mass spectrometry) or binding assays (*e.g*., immunoassays, agglutination assays, and immunochromatographic assays).

30   The method can include measuring or detecting a signal that results from a chemical

reaction, *e.g.*, a change in optical absorbance, a change in fluorescence, the generation of chemiluminescence or electrochemiluminescence, a change in reflectivity, refractive index or light scattering, the accumulation or release of detectable labels from the surface, the oxidation or reduction or redox species, an electrical current or potential,

5   changes in magnetic fields, etc. Suitable detection techniques can detect binding events by measuring the participation of labeled binding reagents through the measurement of the labels via their photoluminescence (*e.g.*, via measurement of fluorescence, time-resolved fluorescence, evanescent wave fluorescence, up-converting phosphors, multi-photon fluorescence, etc.), chemiluminescence, electrochemiluminescence, light

10   scattering, optical absorbance, radioactivity, magnetic fields, enzymatic activity (*e.g.*, by measuring enzyme activity through enzymatic reactions that cause changes in optical absorbance or fluorescence or cause the emission of chemiluminescence). In some examples, detection techniques are used that do not require the use of labels, *e.g.*, techniques based on measuring mass (*e.g.*, surface acoustic wave measurements),

15   refractive index (*e.g.*, surface plasmon resonance measurements), or the inherent luminescence of an analyte, such as a prostate cancer-related marker, for example, AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4.

     In some examples, a binding assay is used. Binding assays for measuring

20   protein levels can use solid phase or homogenous formats. Suitable assay methods include sandwich or competitive binding assays. Examples of sandwich immunoassays are described in U.S. Pat. No. 4,168,146 and U.S. Pat. No. 4,366,241, both incorporated herein by reference. Examples of competitive immunoassays include those disclosed in U.S. Pat. No. 4,235,601, U.S. Pat. No. 4,442,204, and U.S. Pat. No. 5,208,535, all

25   incorporated herein by reference.

     The availability of antibodies specific for prostate cancer-related marker proteins, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins, facilitates the detection and quantitation of these proteins by one of a number of immunoassay methods known in

30   the art, such as those presented in Harlow and Lane (*Antibodies, A Laboratory Manual*,

- 38 -

CSHL, New York, 1988). Methods of constructing such antibodies are known. In addition, antibodies specific for the particular prostate cancer-related marker proteins AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4, for example, are available from several commercial

5   sources. As an example, any of the following antibodies can be used.

For AGR2: #ab82399 and ab209224 from abcam (Cambridge, MA), #PA5-22964 and PA5-34517 from ThermoFisher Scientific (Waltham, MA); #sc-101211 and sc-54561 from Santa Cruz Biotechnology (Dallas, TX); and #AF6068 from R&D Systems (Minneapolis, MN); and monoclonal antibodies P1G4 and P3A5 (Wayner *et al.*, Prostate, **72**:1023-1034, 2012).

10

For AGR3: #ab82400 and ab96345 from abcam; #MA1-10048 and PA5-27222 from ThermoFisher Scientific; #sc-390940 and sc-376653 from Santa Cruz Biotechnology; and #AF6307 and MAB6307 from R&D Systems.

For CRISP3: #ab105951 and ab198048 from abcam, #PA5-45558 and MA5-

15   24001 from ThermoFisher Scientific; #sc-377505 and sc-101378 from Santa Cruz Biotechnology; and #AF2397 and MAB23971 from R&D Systems.

For CCL3: #ab9673 and ab9781 from abcam, #PA1-75246 and PA1-27987 from ThermoFisher Scientific; #sc-1381 and sc-365691 from Santa Cruz Biotechnology; and #AF-270 and MAB450 from R&D Systems.

20   For CEACAM5: #ab190718 and ab91213 from abcam, #MA5-14674 and MIC0101 from ThermoFisher Scientific, #sc-20059 and sc-52390 from Santa Cruz Biotechnology; and #MAB41281 from R&D Systems.

For CEACAM6: #ab78029 and ab134074 from abcam, #MA1-17762 and MA1-17765 from ThermoFisher Scientific, #sc-59899 from Santa Cruz Biotechnology; and

25   #MAB3934 from R&D Systems.

For IL24: #ab182567 and ab115207 from abcam, #PA1-41066 and PA5-21947 from ThermoFisher Scientific, #sc-22769 and sc-8704 from Santa Cruz Biotechnology; and #MAB20571 from R&D Systems.

For MMP9: #ab38898 and ab76003 from abcam, #MA5-15886 and PA5-16851 from ThermoFisher Scientific, #sc-12759 and sc-13520 from Santa Cruz Biotechnology; and #AF1335 and MAB20571 from R&D Systems.

For CXCL14: #ab46010 and ab36622 from abcam, #PA5-28820 and PA5-19645

5   from ThermoFisher Scientific, #sc-130979 from Santa Cruz Biotechnology; and #MAB866 and AF866 from R&D Systems.

For CD90: #ab23894 and ab92574 from abcam, #MA1-35307 and MA1-19029 from ThermoFisher Scientific, #sc-53456 and sc-73163 from Santa Cruz Biotechnology; and #AF2067 and MAB7335 from R&D Systems.

10   For POSTN: #ab14041 and ab92460 from abcam, #P21955 and PA5-34641 from ThermoFisher Scientific, #sc-49479 and sc-67233 from Santa Cruz Biotechnology; and #AF3548 and MAB3548 from R&D Systems.

For SFRP4: #ab154167 and ab32784 from abcam, #PA5-19067 and PA5-29391 from ThermoFisher Scientific, #sc-30152 from Santa Cruz Biotechnology; and

15   #AF1827 from R&D Systems.

Any standard immunoassay format (such as ELISA, Western blot, or RIA assay) can be used to measure protein levels. If desired, a comparison to expression in another sample or comparison to a reference value (or range of values) can be performed. Immunohistochemical techniques can also be utilized for protein detection and

20   quantification. General guidance regarding such techniques can be found in Bancroft and Stevens (*Theory and Practice of Histological Techniques*, Churchill Livingstone, 1982) and Ausubel *et al.* (*Current Protocols in Molecular Biology*, John Wiley & Sons, New York, 1998).

For the purposes of quantitating proteins, a biological sample of the subject that

25   includes cellular proteins (such as urine) can be used. Quantitation of prostate cancer-related marker proteins, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins, can be achieved by immunoassay. The amount of prostate cancer-related marker proteins, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9,

30   CXCL14, CD90, POSTN, and/or SFRP4 proteins, can be assessed in the sample, for

example by contacting the sample with appropriate antibodies (or antibody fragments) specific for each protein, and then detecting a signal (for example present directly or indirectly on the antibody, for example by the use of a labeled secondary antibody).

In one example, an electrochemiluminescence immunoassay is used, such as the V-PLEX[TM] system (Meso Scale Diagnostics, Rockville, MD). In such assays, the primary antibodies for prostate cancer-related marker proteins, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins, (or the corresponding secondary antibodies) are labeled with an electrochemiluminescent label, which allows for ultra-sensitive detection. Examples of electrochemiluminescent labels include, but are not limited to a luminophore (e.g., luminol) on encapsulated gold nanoparticles, a ruthenium pyridine-labeled primary antibody with a biotinylated secondary antibody, and a Sulfo-Tag. Examples of such systems are described in Ge *et al*., *J. Inorg. Oganomet Polym.* 23:1113-21, 2013; US 2014/0072963; US 20100316992, and Sloan *et al*., *Clin. Biochem.* 45:1640-44, 2012, all herein incorporated by reference.

Quantitative spectroscopic approaches methods, such as LC-SRM, LG-SRM [13], PRISM-SRM [14,15], and surface-enhanced laser desorption-ionization (SELDI), can be used to analyze expression of prostate cancer-related marker proteins, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins, in, for example, a urine sample obtained from a subject having or a subject at risk for prostate cancer, such as for intermediate- or high-risk prostate cancer. In some such spectroscopy methods, at least one surrogate peptide for each prostate cancer-related marker protein is measured or detected in the sample (*e.g*., see FIGS. 4 and 15).

In one example, LC-SRM (liquid chromatography-selected reaction monitoring) may be used to detect protein expression for example by using a triple quadrupole spectrometer. Such methods are known in the art (*see, e.g.*, U.S. Pub. No. 2013/0203096). LC-SRM is a liquid chromatography method that can be used for high-throughput selective and sensitive detection of molecules, such as prostate cancer-related proteins, for example, AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6,

IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4. It is useful for quantitating moderately abundant analytes (low µg/mL) in limited sample volumes.

In LC, sample components interact with column packing materials. The flow of a sample through a column is the basis for separation an analyte of interest from the sample. During chromatography, the separation of materials can be effected by variables such as column packing materials, eluent (also known as the mobile phase), elution gradient, and gradient conditions (*e.g.,* temperature). Selected reaction monitoring (SRM) is a method used with tandem mass spectrometry, in which an ion of a particular mass is selected in the first stage of a tandem mass spectrometer, and an ion product of a fragmentation reaction of the precursor ion is selected in the second mass spectrometer stage for detection.

Therefore, in a particular example, the analytes can include prostate cancer-related marker proteins and/or surrogate peptides thereof, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins and/or surrogate peptides thereof. In other examples, the fractionated and pooled analytes consist essentially of or consist of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4 proteins or surrogate peptides thereof or of the combinations of proteins or surrogate peptides listed in FIG. 15. In this context, "consists essentially of" indicates that the fractionated and pooled analytes do not include other prostate cancer-related marker proteins that can be used to accurately predict prostate cancer, but can include other prostate molecules, such as prostate protein expression controls (*e.g*., PSA protein or surrogate peptides thereof).

In another example, LG-SRM (long gradient-selected reaction monitoring) can be used to detect protein expression, for example by using a reversed-phase C18 column and triple quadrupole spectrometer. Such methods are known in the art (see, e.g., Shi *et al.*, Anal Chem., 85(19):9196-9203). LG-SRM is a liquid chromatography method for sensitive quantitation of analytes, such as prostate cancer-related proteins, and can even be used to accurately quantitate low-to-moderately abundant analytes (low µg/mL to high ng/mL), such as CRISP3, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4.

In LG-SRM, a long, shallow LC gradient (*e.g.,* 5 hours compared with a conventional LC protocol that can be about 45 min) using a long LC column is followed by SRM as a second step. The eluting LC peaks containing the target analyte, such as prostate cancer-related proteins or surrogate peptides thereof, for example, AGR2,

5   AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins or surrogate peptides thereof, are, thus, sufficiently separated and resolved for accurate quantitation via SRM.

Therefore, in a particular example, the target analytes include prostate cancer-related marker proteins and/or surrogate peptides thereof, such as AGR2, AGR3,

10   CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins and/or surrogate peptides thereof. In other examples, the target analytes consist essentially of or consist of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4 proteins or surrogate peptides thereof; of the combinations of proteins or surrogate peptides listed in FIG. 15;

15   or of moderate-to-low-abundance proteins or surrogate peptides thereof, such as CRISP3, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4. In this context "consists essentially of" indicates that the target analytes do not include other prostate cancer-related marker proteins that can be used to accurately predict prostate cancer, but can include other prostate molecules, such as prostate protein expression controls (*e.g*., PSA

20   protein or surrogate peptides thereof).

In an additional example, PRISM-SRM (high-pressure, high-resolution separations, intelligent selection, multiplexing-selected reaction monitoring) is used to detect protein expression, for example, by using an ultra-pressure LC (UPLC) system and a triple quadrupole spectrometer. Such methods are known in the art (see, e.g., U.S.

25   Pub. No. 2014/0194304; Shi *et al.*, PNAS, 109(38):15395-15400 (2012); and Shi *et al.*, J Proteome Res., 13(2):875-882 (2014)). PRISM-SRM is a liquid chromatography method for quantitating analyes, such as prostate cancer-related proteins, and can even be used to accurately quantitate low-abundance (low ng/mL to high pg/mL) analytes, such as AGR2, AGR3, CCL3, CEACAM5, and CEACAM6.

In PRISM-SRM, LC-SRM is used as a second step after the target analyte is enriched through a liquid chromatography pre-fractionation step, such as using reverse-phase chromatography. The fractions containing the target analyte, such as prostate cancer-related proteins or surrogate peptides thereof, for example, AGR2, AGR3,

5   CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins or surrogate peptides thereof, can then be pooled. The pooled fractions are enriched in the target analyte(s) and can then undergo a second LC separation step followed by SRM analysis.

Therefore, in a particular example, the fractionated and pooled analytes include

10   prostate cancer-related marker proteins and/or surrogate peptides thereof, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins and/or surrogate peptides thereof. In other examples, the fractionated and pooled analytes consist essentially of or consist of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14,

15   CD90, POSTN, and SFRP4 proteins or surrogate peptides thereof; of the combinations of proteins or surrogate peptides listed in FIG. 15; or of low-abundance proteins or surrogate peptides thereof, such as AGR2, AGR3, CCL3, CEACAM5, and CEACAM6. In this context "consists essentially of" indicates that the fractionated and pooled analytes do not include other prostate cancer-related marker proteins that can be used to

20   accurately predict prostate cancer, but can include other prostate molecules, such as prostate protein expression controls (*e.g.*, PSA protein or surrogate peptides thereof).

In a further example, surface-enhanced laser desorption-ionization time-of-flight (SELDI-TOF) mass spectrometry is used to detect protein expression, for example by using the ProteinChip™ (Ciphergen Biosystems, Palo Alto, CA). Such methods are

25   well-known in the art (for example see U.S. Pat. No. 5,719,060; U.S. Pat. No. 6,897,072; U.S. Pat. No. 6,881,586; Liu AY *et al.*, J Urol, 173:73-78, 2005). SELDI is a solid phase method for desorption in which the analyte is presented to the energy stream on a surface that enhances analyte capture or desorption.

Briefly, one version of SELDI uses a chromatographic surface with a chemistry

30   that selectively captures analytes of interest, such as prostate cancer-related marker

proteins, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins. Chromatographic surfaces can be composed of hydrophobic, hydrophilic, ion exchange, immobilized metal, or other chemistries. For example, the surface chemistry can include binding

5   functionalities based on oxygen-dependent, carbon-dependent, sulfur-dependent, and/or nitrogen-dependent means of covalent or noncovalent immobilization of analytes. The activated surfaces are used to covalently immobilize specific "bait" molecules such as antibodies, receptors, or oligonucleotides often used for biomolecular interaction studies such as protein-protein and protein-DNA interactions.

10   The surface chemistry allows the bound analytes to be retained and unbound materials to be washed away. Subsequently, analytes bound to the surface (such as prostate cancer-related marker proteins) can be desorbed and analyzed by any of several means, for example using mass spectrometry. When the analyte is ionized in the process of desorption, such as in laser desorption/ionization mass spectrometry, the

15   detector can be an ion detector. Mass spectrometers generally include means for determining the time-of-flight of desorbed ions. This information is converted to mass. However, one need not determine the mass of desorbed ions to resolve and detect them: the fact that ionized analytes strike the detector at different times provides detection and resolution of them. Alternatively, the analyte can be detectably labeled (for example

20   with a fluorophore or radioactive isotope). In these cases, the detector can be a fluorescence or radioactivity detector. A plurality of detection means can be implemented in series to fully interrogate the analyte components and function associated with retained molecules at each location in the array.

Therefore, in a particular example, the chromatographic surface includes

25   antibodies that specifically bind prostate cancer-related marker proteins and/or surrogate peptides thereof, such as AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and/or SFRP4 proteins and/or surrogate peptides thereof. In other examples, the chromatographic surface consists essentially of or consists of antibodies that specifically bind AGR2, AGR3, CRISP3,

30   CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4

proteins or surrogate peptides thereof or that specifically bind one of the combinations of proteins or surrogate peptides listed in FIG. 15. In this context "consists essentially of" indicates that the chromatographic surface does not include antibodies that bind any other prostate cancer-related marker proteins that can be used to accurately predict

5   prostate cancer, but can include antibodies that bind other molecules, such as prostate protein expression controls (*e.g*., PSA protein or surrogate peptides thereof).

　　　　In another example, antibodies are immobilized onto the surface using a bacterial Fc binding support. The chromatographic surface is incubated with a sample, such as a urine and/or serum sample obtained from a subject having or a subject at risk

10   for prostate cancer or for intermediate- or high-risk prostate cancer. The antigens present in the sample can recognize the antibodies on the chromatographic surface. The unbound proteins and mass spectrometric interfering compounds are washed away and the proteins that are retained on the chromatographic surface are analyzed and detected by SELDI-TOF. The MS profile from the sample can be then compared using

15   differential protein expression mapping, whereby relative expression levels of proteins at specific molecular weights are compared by a variety of statistical techniques and bioinformatic software systems.

Administration of Therapy

20   　　　　Subjects analyzed with the disclosed methods and who are found to have prostate cancer can be selected for treatment. For example, subjects with prostate cancer or with intermediate-to-high-risk prostate cancer found to have increased expression of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4 can be administered therapy for prostate cancer.

25   Currently, the standard of care for prostate cancer can vary, but level of risk can be a factor. For example, a subject may be found to have low-risk prostate cancer, such as a patient with increased levels of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4 but with MMP9 levels expected from a patient with low-risk prostate cancer. In some examples,

30   subjects with low-risk prostate cancer may be treated using watchful waiting or active

surveillance, both of which entail monitoring the cancer for changes and the subject for symptoms. Given that more invasive treatments entail a greater potential for side effects, studies suggest that active surveillance is the best choice for patients with low-risk prostate cancer.

5   In other examples, surgical removal of the prostate can be a treatment for low-risk prostate cancer or prostate cancers that do not respond to radiation therapy. In additional examples, subjects with any stage of prostate cancer, such as subjects with increased expression of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4 can be treated with radiation therapy,

10   such as using ionizing radiation to kill prostate cancer cells. In some other examples, subjects with either low- or intermediate-risk prostate cancer can be treated using brachytherapy, for example, where small radioactive particles, such as iodine-125 or palladium-103, are directly injected into the tumor, providing localized X-rays at the site of the tumor. In additional examples, ultrasound, such as high-intensity focused

15   ultrasound (HIFU) is used where a subject has a recurrent case of prostate cancer, such as where a subject was successfully treated for prostate cancer but subsequently had increased expression of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4.

          In further examples, a subject can be treated with hormone therapy, such as by

20   modulating the levels of testosterone in the body, where the subject has either recurrent prostate cancer, for example, a subject that was successfully treated for prostate cancer but subsequently had increased expression of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4, or a subject that has high-risk prostate cancer, for example, a subject that has increased

25   expression of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4 and increased expression of MMP9 compared with the expression expected from a patient with low-risk prostate cancer.

          In some examples, at least a portion of the prostate cancer is surgically removed (for example via cryotherapy), irradiated, chemically treated (for example *via*

30   chemoembolization), or combinations thereof, as part of the therapy. For example, a

subject having prostate cancer can have all or part of the tumor surgically excised prior to administration of additional therapy.

Exemplary agents that can be used include one or more anti-neoplastic agents, such as radiation therapy, chemotherapeutic, biologic (e.g., immunotherapy), and anti-

5   angiogenic agents or therapies. Methods and therapeutic dosages of such agents are known to those skilled in the art, and can be determined by a skilled clinician. These therapeutic agents (which are administered in therapeutically effective amounts) and treatments can be used alone or in combination. In some examples, 1, 2, 3, 4 or 5 different anti-neoplastic agents are used as part of the therapy.

10   In one example the therapy includes administration of one or more chemotherapy immunosuppressants (such as Rituximab, steroids) or cytokines (such as GM-CSF). Chemotherapeutic agents are known (see for example, Slapak and Kufe, Principles of Cancer Therapy, Chapter 86 in Harrison's Principles of Internal Medicine, 14th edition; Perry et al., Chemotherapy, Ch. 17 in Abeloff, Clinical Oncology 2nd ed.,

15   2000 Churchill Livingstone, Inc; Baltzer and Berkery. (eds): Oncology Pocket Guide to Chemotherapy, 2nd ed. St. Louis, Mosby-Year Book, 1995; Fischer Knobf, and Durivage (eds): The Cancer Chemotherapy Handbook, 4th ed. St. Louis, Mosby-Year Book, 1993). Exemplary chemotherapeutic agents that can be used with the therapy include but are not limited to, carboplatin, cisplatin, paclitaxel, docetaxel, doxorubicin,

20   epirubicin, cabaziatxel, estramustine, vinblastine, topotecan, irinotecan, gemcitabine, iazofurine, etoposide, vinorelbine, tamoxifen, valspodar, cyclophosphamide, methotrexate, fluorouracil, mitoxantrone, and Doxil® (liposome encapsulated doxiorubicine). In one example the therapy includes docetaxel and prednisone. In one example the therapy includes cabaziatxel.

25   In one example, the therapy includes administering one or more of a microtubule binding agent, DNA intercalator or cross-linker, DNA synthesis inhibitor, DNA and/or RNA transcription inhibitor, antibodies, enzymes, enzyme inhibitors, and gene regulators.

Microtubule binding agents interact with tubulin to stabilize or destabilize

30   microtubule formation thereby inhibiting cell division. Examples of microtubule

binding agents that can be used as part of the therapy include, without limitation, paclitaxel, docetaxel, vinblastine, vindesine, vinorelbine (navelbine), the epothilones, colchicine, dolastatin 10, nocodazole, and rhizoxin. Analogs and derivatives of such compounds also can be used. For example, suitable epothilones and epothilone analogs are described in International Publication No. WO 2004/018478. Taxoids, such as paclitaxel and docetaxel, as well as the analogs of paclitaxel taught by U.S. Patent Nos. 6,610,860; 5,530,020; and 5,912,264 can be used.

The following classes of compounds can be used as part of the therapy: suitable DNA and/or RNA transcription regulators, including, without limitation, anthracycline family members (for example, daunorubicin, doxorubicin, epirubicin, idarubicin, mitoxantrone, and valrubicin) and actinomycin D, as well as derivatives and analogs thereof. DNA intercalators and cross-linking agents that can be administered to a subject include, without limitation, platinum compounds (for example, carboplatin, cisplatin, oxaliplatin, and BBR3464), mitomycins, such as mitomycin C, bleomycin, chlorambucil, cyclophosphamide, as well as busulfan, dacarbazine, estramustine, and temozolomide and derivatives and analogs thereof. DNA synthesis inhibitors suitable for use as therapeutic agents include, without limitation, methotrexate, 5-fluoro-5'-deoxyuridine, 5-fluorouracil and analogs thereof. Examples of suitable enzyme inhibitors include, without limitation, camptothecin, etoposide, exemestane, trichostatin and derivatives and analogs thereof. Suitable compounds that affect gene regulation include agents that result in increased or decreased expression of one or more genes, such as raloxifene, 5-azacytidine, 5-aza-2'-deoxycytidine, tamoxifen, 4-hydroxytamoxifen, mifepristone, and derivatives and analogs thereof. Kinase inhibitors include imatinib, gefitinib, and erolitinib that prevent phosphorylation and activation of growth factors.

In one example, the therapy includes folic acid (for example, methotrexate and pemetrexed), purine (for example, cladribine, clofarabine, and fludarabine), pyrimidine (for example, capecitabine), cytarabine, fluorouracil, gemcitabine, and derivatives and analogs thereof. In one example, the therapy includes a plant alkaloid, such as podophyllum (for example, etoposide) and derivatives and analogs thereof. In one

example, the therapy includes an antimetabolite, such as cytotoxic/antitumor antibiotics, bleomycin, hydroxyurea, mitomycin, and derivatives and analogs thereof. In one example, the therapy includes a topoisomerase inhibitor, such as a topoisomerase I inhibitor (*e.g*., topotecan, irinotecan, indotecan, indimitecan, camptothecin and

5 lamellarin D) or a topoisomerase II inhibitor (*e.g*., etoposide, doxorubicin, daunorubicin, mitoxantrone, amsacrine, ellipticines, aurintricarboxylic acid, ICRF-193, genistein, and HU-331), and derivatives and analogs thereof. In one example, the therapy includes a photosensitizer, such as aminolevulinic acid, methyl aminolevulinate, porfimer sodium, verteporfin, and derivatives and analogs thereof. In one example, the

10 therapy includes a nitrogen mustard (for example, chlorambucil, estramustine, cyclophosphamide, ifosfamide, and melphalan) or nitrosourea (for example, carmustine, lomustine, and streptozocin), and derivatives and analogs thereof.

Other therapeutic agents, for example anti-tumor agents, that may or may not fall under one or more of the classifications above, also are suitable for therapy. By

15 way of example, such agents include adriamycin, apigenin, rapamycin, zebularine, cimetidine, amsacrine, anagrelide, arsenic trioxide, axitinib, bexarotene, bevacizumab, bortezomib, celecoxib, estramustine, hydroxycarbamide, lapatinib, pazopanib, masoprocol, mitotane, tamoxifen, sorafenib, sunitinib, vandetanib, tretinoin, and derivatives and analogs thereof.

20 In one example, the therapy includes one or more biologics, such as a therapeutic antibody, such as monoclonal antibodies. Examples of such biologics that can be used include one or more of bevacizumab, cetuximab, panitumumab, pertuzumab, trastuzumab, bevacizumab (Avastin®), ramucirumab, and the like. In specific examples, the antibody or small molecules used as part of the therapy include

25 one or more of the monoclonal antibodies cetuximab, panitumumab, pertuzumab, trastuzumab, bevacizumab (Avastin®), ramucirumab, or a small molecule inhibitor such as gefitinib, erlotinib, and lapatinib.

In some examples the therapy includes administration of one or more immunotherapies, which may include the biologics listed herein. In specific examples,

30 the immunotherapy includes therapeutic cancer vaccines, such as those that target PSA

(*e.g.*, ADXS31-142), prostatic acid phosphatase (PAP) antigen, TARP, telomerase (e.g., GX301) or that deliver 5T4 (e.g.., ChAdOx1 and MVA); antigens NY-ESO-1 and MUC1; antigens hTERT and survivin; prostate-specific antigen (PSA) and costimulatory molecules (*e.g.*, LFA-3, ICAM-1, and B7.1) directly to cancer cells, such

5   as rilimogene galvacirepvac. Other examples of therapeutic vaccines include DCVAC, sipuleucel-T, pTVG-HP DNA vaccine, pTVG-HP, JNJ-64041809, PF-06755992, PF-06755990, and pTVG-AR. In other examples, the immunotherapy includes oncolytic virus therapy, such as aglatimagene besadenovec, HSV-tk, and valacyclovir. In additional examples, the immunotherapy can include checkpoint inhibitors, such as

10   those that target PD-1 (*e.g.*, nivolumab, pembrolizumab, durvalumab, and atezolizumab), CTLA-4 (*e.g.*, tremelimumab and ipilimumab), B7-H3 (*e.g.*, MGA271), and CD27 (e.g., CDX-1127). The protein MGD009 may also be used in another example. In specific examples, the immunotherapy can also include adoptive cell therapy, such as those that include T cells engineered to target NY-ESO-1 and those

15   that include natural killer (NK) cells. In some examples, the immunotherapy can include adjuvant immunotherapies, such as sipuleucel-T, indoximod, and mobilan. In other specific examples, the immunotherapy includes one or more of tisotumab vedotin, sacituzumab govitecan, LY3022855, BI 836845, vandortuzumab vedotin, and BAY2010112, and MOR209/ES414. In additional examples, the immunotherapy can

20   include cytokines, such as CYT107, AM0010, and IL-12.

In some examples, the subject receiving the therapy is also administered interleukin-2 (IL-2), as part of the therapy, for example *via* intravenous administration. In particular examples, IL-2 is administered at a dose of at least 500,000 IU/kg as an intravenous bolus over a 15 minute period every eight hours beginning on the day after

25   administration of the peptides and continuing for up to 5 days. Doses can be skipped depending on subject tolerance.

In some examples, the subject receiving the therapy is also administered a fully human antibody to cytotoxic T-lymphocyte antigen-4 (anti–CTLA-4) as part of the therapy, for example *via* intravenous administration. In some example subjects receive

at least 1 mg/kg anti-CTLA-4 (such as 3 mg/kg every 3 weeks or 3 mg/kg as the initial dose with subsequent doses reduced to 1 mg/kg every 3 weeks).

In one specific example for a subject with prostate cancer, such as a subject with increased expression of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4, the therapy can include one or more of abiraterone acetate, bicalutamide, cabazitaxel, casodex (bicalutamide), degarelix, docetaxel, enzalutamide, flutamide, goserelin acetate, jevtana (cabazitaxel), leuprolide acetate, lupron (leuprolide acetate), lupron depot (leuprolide acetate), lupron depot-3 month (leuprolide acetate), lupron depot-4 month (leuprolide acetate), lupron depot-ped (leuprolide acetate), mitoxantrone hydrochloride, nilandron (nilutamide), nilutamide, provenge (sipuleucel-t), radium 223 dichloride, sipuleucel-T, taxotere (docetaxel), viadur (leuprolide acetate), xofigo (radium 223 dichloride), xtandi (enzalutamide), zoladex (goserelin acetate), and zytiga (abiraterone acetate).

In another specific example for a subject with prostate cancer, such as a subject with increased expression of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4, the therapy can include one or more of chemotherapy drugs, such as cabazataxel (Jevtana®), docetaxel (Taxotere®), mitoxantrone (Teva®), or androgen deprivation therapy (ADT), such as with abiraterone Acetate (Zytiga®), bicalutamide (Casodex®), buserelin Acetate (Suprefact®), cyproterone Acetate (Androcur®), degarelix Acetate (Firmagon®), enzalutamide (Xtandi®), flutamide (Euflex®), goserelin Acetate (Zoladex®), histrelin Acetate (Vantas®), leuprolide Acetate (Lupron®, Eligard®), triptorelin Pamoate (Trelstar®). The therapy can also include drugs to treat bone metastases (bisphosphate therapy), such as alendronate (Fosamax®), denosumab (Xgeva®), pamidronate (Aredia®), zoledronic acid (Zometa®), or radiopharmaceuticals, such as radium 223 (Xofigo®), strontium-89 (Metastron®), and samarium-153 (Quadramet®).

The therapy can be administered in cycles (such as 1 to 6 cycles), with a period of treatment (usually 1 to 3 days) followed by a rest period. But some therapies can be administered every day.

The disclosure is illustrated by the following non-limiting Examples.

EXAMPLES

Example 1

5   **Methods and Materials**

This example provides technical details and procedures, including relevant instrument settings and materials, used to obtain the protein expression data from patient urine samples discussed in the Examples below.

10   *Urine collection, processing, and protein digestion*

Human urine samples were approved by an Institutional Review Boards with written consent of donors, and the samples were anonymized before being given to researchers. The suffix N was added to the sample codes to denote non-cancer, and the suffix C to denote cancer from pre-operative (pre-op) patients. Urine from post-

15   operative patients (post-op) was collected after surgical resection of the prostate. All urine samples were freshly collected for this study; no archived samples were used.

Collected voided urine samples were processed within 2 h (to isolate RNA as well). The samples were centrifuged at 1,200 rpm for 5 min, and the supernatant was stored at -80°C. Fifteen to 90 mL of urine was desalted and concentrated using

20   Amicon® Ultra-15 (3 kDa molecular weight cut-off, Millipore, USA) [10]. Protein concentrations were determined by the BCA assay (Pierce, USA). Concentrated urinary proteins from each sample, ranging from 200 to 300 μg, were denatured and reduced with 8 M urea and 10 mM DTT in 50 mM NH_4HCO_3, pH 8.0 for 1 h at 37°C. Protein cysteine residues were alkylated with 40 mM iodoacetamide for 1 h at room

25   temperature in the dark. The resulting sample was diluted 6-fold with 50 mM NH_4HCO_3, pH 8.0, and digested by sequencing-grade modified porcine trypsin (Promega, USA) at 1:50 trypsin:protein (w/w) overnight at 37°C. The resulting digest was desalted by using a 1 mL SPE C18 column (Supelco, USA) as described previously [9]. The final tryptic peptide concentration was determined by the BCA assay. The

30   peptide sample was diluted to 0.5 μg/μL with 0.1% formic acid in water, and heavy

isotope-labeled synthetic peptides were spiked in at an equimolar concentration, 10

fmol/µL of crude heavy peptides for the secreted protein markers, 10 fmol/µL of pure

heavy peptide IVGGWECcamEK (Ccam: cysteine residue synthesized as

carbamidomethyl cysteine), and 1 fmol/µL of pure heavy peptide LSEPAELTDAVK

5   for PSA, due to the response difference between the two PSA peptides.

Database query

The human urine proteome databases archived in PeptideAtlas

(www.peptideatlas.org) were queried for data entries of marker identifiers. The

10   UrinePA build contained high confidence peptide and protein identifications obtained

from five labs using tandem mass spectrometry proteomics [16]. About 2,500 non-

redundant proteins were cataloged at a 1% false discovery rate. Another database listed

587 entries of a "Core Urinary Proteome", which was established from an in-depth

analysis of second morning urine obtained over three days from seven healthy

15   volunteers between 25-35 years old [17].

SRM assay configuration

Urea, dithiothreitol (DTT), iodoacetamide, ammonium formate, trifluoroacetic

acid (TFA), and formic acid were purchased from Sigma (USA). The synthetic

20   peptides labeled with $^{13}C/^{15}N$ on C-terminal lysine and arginine residues were from

Thermo Scientific (USA). The heavy peptides for PSA protein were estimated to be of

>95% purity by HPLC.

Ten surrogate peptides were first chosen for the protein markers based on in

silico trypsin digestion and existing MS/MS data, the Global Proteome Machine

25   (GPM), and PeptideAtlas. These peptides were then evaluated by ESP predictor [18]

and CONSeQuence [19] software. Three to five peptides with moderate hydrophobicity

and high scores from the prediction tools were selected for peptide synthesis. The

synthesized crude heavy-isotope labeled peptides were further evaluated for peptide

response and fragmentation pattern. Optimal collision energy (CE) values were

30   achieved by direct infusion of the individual peptides and/or multiple LC-SRM runs

with CE ramping. The best performing peptides were used for detection and
quantification of the secreted protein markers. For each peptide, the three best
transitions and matrix interference were determined. The relative intensity ratios among
the three selected transitions for SRM were predefined by the internal standard heavy

5   peptides in buffer. The matrix interference for a given transition that fell into mass
widths Q1 and Q3 from co-eluting peptides was identified by a deviation from the
expected relative intensity ratios among the transitions. The transition with no matrix
interference was used for marker quantification in urine samples.

10   *LG-SRM measurement*

        The LG-SRM approach was previously demonstrated to enable reproducible
quantification of target proteins at ~10 ng/mL levels in nondepleted human serum [13].
Typically, 4 µL of the tryptic digest samples with a peptide concentration of 0.5 µg/µL
was directly loaded onto a capillary reversed-phase column, 75 µm inner diameter (i.d.)

15   × 150 cm length, packed in-house with 3-µm Jupiter C18 bonded particles
(Phenomenex, USA) to permit long gradient separation without a trap column with its
dead volume affecting peptide retention time. Peptide separations were performed at a
mobile phase flow rate of 100 nL/min on a binary pump system using 0.1% formic acid
in water as phase A and 0.1% formic acid in 90% acetonitrile as phase B. The profile

20   for a 300-min gradient time was 5–15% B in 27 min, 15–25% B in 140 min, 25–35% B
in 73 min, and 35–90% B in 60 min. The TSQ Vantage mass spectrometer was
operated in the manner as previously described [13].

     PRISM-SRM measurement

25   The PRISM-SRM approach has been previously described for quantification of
low-abundance proteins in human plasma or serum [14]. Briefly, high-resolution
reversed-phase capillary LC with pH 10 mobile phase was used as the first dimensional
separation of peptides from trypsin-digested human urine proteins. Following
separation, the column eluent was automatically collected every minute into a 96-well

30   plate during a ~100-min LC run while on-line SRM monitoring of heavy internal

standard peptides was performed on a small split stream of the flow. Intelligent

selection (termed *i*Selection) of target peptide fractions was achieved based on the on-

line SRM signal of internal standard peptides. Prior to peptide fraction collection, 17

µL of water was added to each well to minimize excessive loss of peptides and dilute

5   the peptide fractions (~1:7) for LC-SRM analysis.

         Following iSelection, the target peptide-containing fractions were subjected to

LC-SRM measurement. All peptide fractions were analyzed by using the

nanoACQUITY UPLC® system (Waters Corporation, USA) coupled on-line to a TSQ

Vantage triple quadrupole mass spectrometer (Thermo Scientific). Solvents used were

10   0.1% formic acid in water (mobile phase A) and 0.1% formic acid in 90% acetonitrile

(mobile phase B). Peptide separations were performed at a mobile phase flow rate of

400 nL/min using an ACQUITY UPLC BEH 1.7 µm C18 column (75 µm i.d. × 10 cm),

which was connected to a chemically etched 20 µm i.d. fused-silica emitter via a Valco

stainless steel union. Four microliters of individual peptide fractions (total volume 20

15   µL) following PRISM were injected for LC separation using a binary gradient of 10-

20% phase B in 7 min, 20-25% phase B in 17 min, 25-40% phase B in 1.5 min, 40-95%

phase B in 2.5 min, and 95% phase B in 6 min for a total time of ~35 min. The TSQ

Vantage was operated in the same manner as previously described [9,14]. A scan width

of 0.002 *m/z* and a dwell time of 40 ms were set for all SRM transitions.

20

Proteomics data analysis

         SRM data were analyzed using Skyline software [20]. Peak detection and

integration were determined based on two criteria: (1) the same retention time and (2)

approximately the same relative SRM peak intensity ratios across multiple transitions

25   between light (L) peptides and heavy (H) peptide standards [9,14,21]. All data were

manually inspected to ensure correct peak detection and accurate integration. The

signal to noise ratio (S/N) was calculated by the peak apex intensity over the highest

background noise in a retention time region of ±15 s for the target peptides [14,21].

The background noise levels were conservatively estimated by visually inspecting

30   chromatographic peak regions. Quantifiable endogenous surrogate peptides should

have SRM signals with S/N ≥10. The RAW data from the TSQ Vantage were loaded into Skyline to create high resolution figures of extracted ion chromatograms (XICs) of multiple transitions monitored for the target peptides/proteins. GraphPad Prism 6.0 was used for statistical analysis and plotting; *P* <0.05 was considered statistically significant

5   [9]. Receiver operating characteristic (ROC) curves were produced in terms of sensitivity and specificity of protein markers at their specific cutoff values. The optimal cutoff was the point with the best sum of sensitivity and specificity. Multivariate evaluative analysis for various combinations of protein markers was performed using SPSS 16 by logistic regression to find the best-fitting model for each comparison group.

10

AGR2 ELISA

Urinary AGR2 was measured by a developed sandwich ELISA based on monoclonal antibodies P1G4 (IgG1) to capture and P3A5 (IgG2a) to detect as described previously [8]. Recombinant AGR2 (GenWay Biotech, San Diego, CA) was used to

15   generate a standard curve. Purified P1G4 at 1:1,000 was used to capture the analyte, and purified P3A5 at 1:1,000 was used for detection followed by HRP-conjugated anti-mouse IgG2a (SouthernBiotech, USA). The chromogenic reaction was recorded at plate reader setting of λ= 405 nm. Culture media of the AGR2-secreting prostate cancer cell line PC3 was used as a positive control. Buffer was the negative control.

20

Calculations

Secreted protein concentration calculation in urine (ng/100 µg of total urinary protein):

[secreted protein] = L/H × IS (fmol/µL) × MW × 10^{-6} (ng/fmol) × 200 µL

IS: internal standard

25   MW: protein molecular weight (g/mol)

100 µg of urinary protein: 200 µL of 0.5 µg/ µL urine peptide

Example 2

Tumor-associated secreted proteins in human urine

This example describes methods that were used to determine the tumor-associated secreted proteins present in urine.

5   Through comparison of cell type-specific transcriptomes, genes showing elevated tumor expression and encoding secreted/extracellular proteins were identified from both the epithelial and stromal compartments. Furthermore, gene expression analysis indicated that many showed differential expression among tumors of different Gleason scores. The epithelial-derived marker candidates included AGR2, AGR3,

10   CRISP3, CEACAM5, CEACAM6, CCL3, CCL4, IL24, and MMP9; the stromal-derived candidates included CXCL14, CD90, IL24, MMP9, POSTN, SFRP4, and WISP1. The database PeptideAtlas was interrogated to determine if these proteins were present in non-cancer urine. The query result is shown in FIG. 1. In the UrinePA archive of available mass spectrometry datasets, the "observed" qualifier was used to

15   indicate protein abundance. UMOD and ALB were the most abundant with the observed times 24,115 and 33,149, respectively. Other proteins, such as APOD (2,352), LMAN2 (639), AMBP (15,223), F2 (1,164), PTGDS (4,479), and AZGP1 (1,903), were the major urinary proteins that could be visualized by staining of a polyacrylamide gel electrophoregram [10]. Prostatic proteins, such as KLK3/PSA (193), KLK2 (2), ACPP

20   (415), MSMB (21), PSCA (123), and PRSS8 (230), were also cataloged as was stromal cell PENK (6) [22]. ERG (0), a cytoplasmic protein in prostate cancer cells and endothelial cells [23], was not found. Non-secreted UPK3A (3) was present for bladder cells. A group of cell surface CD antigens expressed by luminal and cancer cells (CD26, CD10, CD57, CD38, CD107a, and CD107b) were also found. CD45 (0) was

25   queried for leukocytes, and CD31 (9) for endothelial cells. Of the marker candidates, CRISP3 (65), CEACAM5 (21), CEACAM6 (5), THY1/CD90 (261), MMP9 (115), and SFRP4 (17) were cataloged in non-cancer, demonstrating that they could be detected in all urine samples. Therefore, those proteins that were not encountered in the database of healthy donors were more likely derived from diseased tissues, such as prostate

30   cancer in the urinary system.

Example 3

Development of sensitive multiplexed peptide assays for prostate cancer-associated protein markers

5   This example describes methods that were used to develop sensitive multiplexed peptide assays for prostate cancer-associated protein markers.

Selection of the most suitable surrogate peptides is important for sensitive accurate quantification of target proteins in patient specimens. For each prostate cancer protein marker, the surrogate peptides selected for assay development are listed in FIG.

10   2. For PSA, IVGGWECcamEK and LSEPAELTDAVK were demonstrated to be most effective for quantification [14, 24]. For the others, a pooled prostate cancer patient urine sample was used to configure the final SRM assays with an evaluation of matrix interference, endogenous peptide detectability, and the peptide SRM response. LG-SRM was used first to simultaneously measure the 14 candidates due to its moderate

15   sensitivity (≥10-fold higher than LC-SRM) and higher multiplexing capability (~three times higher than LC-SRM). Seven proteins, CD90, CRISP3, CXCL14, IL24, MMP9, POSTN, and SFRP4 (note protein names of these genes, which could be different, are given in the figures), were confidently detected and quantified by at least one surrogate peptide (FIGS. 3A-3B). The ultrasensitive PRISM-SRM (≥ 20-fold higher in sensitivity

20   than LG-SRM) was then used to measure the other seven candidates. Five proteins, AGR2, AGR3, CCL3, CEACAM5, and CEACAM6, were reliably detected and quantified (FIGS. 3A-3B). Based on the LG-SRM and PRISM-SRM results, SRM assays were established for each of the detectable peptides with the three best transitions without matrix interference, and the best transition was used for

25   quantification (FIG. 4). From the assay results, the 12 detected markers were grouped into seven moderate-to-low abundance proteins for LG-SRM analysis and five low abundance proteins for PRISM-SRM analysis. The two undetected proteins, CCL4 and WISP1, were excluded from further testing. The configured SRM assays were used for measuring the 12 protein biomarkers in a cohort of 14 cancer (pre-op) and 6 non-cancer

(healthy control) urine samples. Another cohort of post-op urine samples was collected and analyzed separately to determine the origin of urinary PSA.

Example 4

Quantification accuracy of individual surrogate peptides

This example shows assessment of the quantification accuracy for individual surrogate peptides to determine which surrogate peptides are suitable for diagnosis.

Multiple surrogate peptides were tested for quantification of a specific protein in urine. Without posttranslational modifications or undocumented amino acid changes in the surrogate peptides, their measured concentrations should have a high degree of correlation across all samples because the surrogate peptide level was stoichiometric to that of their cognate protein [25]. With any changes in the peptide sequence, the level of the unmodified surrogate peptides would be lower, affecting accurate measurement of their corresponding proteins. For these differences, each surrogate peptide could potentially represent a distinctive signature with diagnostic value [26].

To evaluate the quantification accuracy of surrogate peptides, a correlation analysis of the L/H ratios between the surrogate peptides from the same protein was carried out. MMP9 was represented by four quantifiable surrogate peptides, and the Pearson correlation coefficients ranged from 0.59 between FQTFEGDLK and SLGPALLLLQK to 0.93 between AVIDDAFAR and FQTFEGDLK, which suggested that multiple MMP9 isoforms could exist in these clinical urine samples (FIGS. 5A-5F). For CD90, low correlation coefficients between VTSLTACLVDQSLR and the other two peptides were obtained, whereas a good correlation, $R^2 = 0.72$, was obtained with the other two peptides (FIGS. 6A-6B). These data show that unknown modifications were present in the VTSLTACLVDQSLR sequence in several urine samples, making this peptide unsuitable for quantitative measurement of CD90. However, such modifications could be cancer-specific.

A high correlation between the two widely used PSA peptides, IVGGWECcamEK and LSEPAELTDAVK, was obtained with $R^2 = 0.99$ across the 20 urine samples (FIGS. 7A-7B). Closer examination revealed three data points from urine

samples P08-015C, P07-040C, and P06-017C that deviated from the correlation plot (FIGS. 7A-7B and FIG. 8). These data show that the two PSA peptides in these urine samples contained amino acid alterations. For example, the P06-017C urine showed a lower L/H ratio for LSEPAELTDAVK. The experimental results indicate an amino

5   acid change in the LSEPAELTDAVK sequence (FIG. 9). These data are supported by a recent discovery of PSA proteoforms, a nonsynonymous mutation L132I (rs2003783) within LSEPAE(L/I)TDAVK [27]. For the other two urine samples with the lower L/H ratios for IVGGWECcamEK, a similar type of change or others could be present [28].

10   **Example 5**

                    Urinary PSA exclusively secreted by the prostate gland

                    This example shows that urinary PSA is primarily secreted by the prostate gland and that its secretion from other sources is negligible.

                    To quantitatively evaluate the percentage of urinary PSA originating from the

15   prostate gland, LC-SRM was used to measure its concentration in seven post-op urine samples (i.e., from men without a prostate). The measured PSA levels ranged from 0.02 ng/100 µg to 2.95 ng/100 µg of total urinary protein with an average value of 0.98 ng/100 µg (and a median value of 0.41 ng/100 µg, FIG. 10). When compared to the PSA levels in the non-cancer and cancer samples with an average value of 110.89

20   ng/100 µg of total urinary protein (and a median value of 28.68 ng/100 µg), the percentage of PSA in the post-op urine was only ~1% (~1.5% median value, FIG. 10). In non-cancer and cancer, urinary PSA can be contributed by all possible sources, whereas in post-op, it cannot be contributed from an absent prostate gland. Thus, the data show that urinary PSA is secreted mostly from the prostate gland, and contribution

25   from other sources in the urinary system is negligible.

                                            Example 6

           Normalizing the concentration of cancer-associated proteins secreted using PSA

                    This example shows normalization of cancer-associated secreted protein

30   concentration using PSA to facilitate differentiation of cancer and non-cancer samples.

In the SRM targeted measurement, the reported L/H peak area ratios of surrogate peptides were proportional to the concentrations of their cognate protein and are expressed as ng/100 µg of total urinary protein because the same amount of the heavy internal standards was added to the digested peptide mixtures for all 20 samples,

5   which have the same peptide concentrations. Thus, the L/H ratio is the adjusted concentration of the target protein in urine (against the total amount of the urinary proteins [9], FIGS. 11A-11B). This adjustment accounts for substantial variations in urinary protein concentration among donors and sometimes among donations from the same donor. For most surrogate peptides measured, the cancer urine showed higher

10   median L/H values than the non-cancer urine; while for several other peptides (CRISP3, CXCL14, IL24, and SFRP4), a lower or equal median L/H value in cancer vs. non-cancer was found. A Mann-Whitney U test of the surrogate peptide L/H ratios revealed no significant difference between cancer and non-cancer for all of the secreted protein markers (FIG. 12). Because the secreted urinary proteome is a pool comprised of all

15   proteins produced by organs along the urinary tract (see PeptideAtlas query result), "normalization" to the secreted proteins solely produced by the prostate was necessary. The post-op urine analysis showed that PSA levels, which were proportional to the size of the gland and the number of secretory cells, luminal and cancer, could be used for this normalization. PSA normalization was used to present urinary AGR2 levels as

20   AGR2/PSA ratios [9]. The amount of PSA was similarly used in the urine PCA3 assay in which the urinary PCA3 score was generated by normalizing the PCA3 transcript levels to those of PSA transcript [29].

Based on the above peptide correlation analysis, both PSA peptides could be used to determine PSA protein concentrations in the cohort. The secreted protein

25   marker/PSA concentration ratios were obtained by dividing the L/H peak area ratios of surrogate peptides for the protein markers by the L/H peak area ratio of the PSA peptide IVGGWECcamEK (FIGS. 13A-13B). After PSA normalization, a significant difference between the cancer and non-cancer urine was observed for the marker peptides [except for LPQTLSR of AGR2 (see below), VTSLTACLVDQSLR of CD90,

30   and MVIITTK of CXCL14] with the *P* values ranging from 0.015 to 0.035 (FIG. 12 and

FIGS. 14A-14F). These results demonstrate the utility of PSA protein normalization in prostate cancer urine biomarker performance.

ROC analyses with 95% confidence intervals show that the peptides with P <0.05 produced AUC values >0.80, while for the three peptides with P >0.05, the AUC

5   values produced were <0.80 (FIG. 12 and FIGS. 14A-14B). With PSA normalization, all the peptide signatures showed values of the sum of sensitivity and specificity ranging from 1.60 to 1.79 at the optimal cutoff points. These data show good biomarker performance in differentiation of prostate cancer from non-cancer through urine analysis. In addition, surrogate peptides with a good correlation produced the same

10   AUC values. For example, the AUC values obtained from the four MMP9 surrogate peptides, AVIDDAFAR, FQTFEGDLK, LGLGADVAQVTGALR, and SLGPALLLLQK, were 0.82, 0.86, 0.86, and 0.86, respectively. The AUC values from the two well-correlated CD90 surrogate peptides, VLYLSAFTSK and HVLFGTVGVPEHTYR, were 0.86 and 0.87, respectively. The VTSLTACLVDQSLR

15   peptide without significant correlations produced an AUC value of 0.77 (FIG. 12). Multi-marker performance was also assessed by using a multivariate analysis of various combinations of the surrogate peptides (FIG. 15). The best combination consisted of LPQTLSR/AGR2, LYTYEPR/AGR3, SDLVNEEATGQFR/CEACAM5, VTSLTACLVDQSLR/CD90, and GVCISPEAIVTDLPEDVK/SFRP4. With a P value

20   of 0.002 and an AUC value of 0.95, this peptide grouping outperformed any single marker (FIGS. 14E-14F, FIG. 12, and FIG. 15).

Example 7

Transient increase in urinary AGR2 in non-cancer urine

25   This example demonstrates transient increases in urinary AGR2 in non-cancer urine and a persistently high urinary AGR2 in cancer urine.

A fraction of non-cancer urine was found to have an above-background amount of urinary AGR2 [30], although urine proteome database query yielded no AGR2 identifier (*cf*. FIG. 1). The AGR2 levels in a middle-age non-cancer male (with normal

30   serum PSA (sPSA)) over a two-week period are shown in FIG. 16B. The levels

increased on day 4 (for both morning and afternoon urine) followed by a decrease to the baseline level on subsequent days. Treatment of urine samples by the addition of alcohol (to lyse shed AGR2$^+$ bladder cells [31]) did not raise the level of detectable urinary AGR2 (bars 3 vs. 4, FIG. 16B). Thus, some day-to-day physiological

5   differences could account for AGR2 in the urine of non-cancer samples. This phenomenon affected the AUC for AGR2. Unlike non-cancer samples, AGR2 in cancer urine samples remained above the background, and the urine contained other marker proteins. Testing this hypothesis required a suitable patient donor to volunteer donations every day over a similar period of two weeks. Nevertheless, the AUC value

10   obtained for AGR2 was 0.74 (FIG. 16A) by ELISA and 0.77 by PRISM-SRM. Hence, both biomarker measurement methods produced concordant results.

Example 8

Detection of clinically significant cancers by secreted protein markers

15   This example compares differentiative power in identifying the low-risk from the high-risk cancers among secreted protein markers. This example also shows that MMP9 expression can be used to differentiate low- and high-risk prostate cancer.

        The prostate cancer cohort was grouped into either low volume/low grade: Gleason score ≤6 and dominant tumor volume ≤0.5 cc [31] or clinically significant: not

20   meeting the criteria for the low volume/low grade disease. The marker levels for low-risk cancer and high-risk cancer were analyzed by Mann-Whitney's U test (FIGS. 17 and 18A-18B). For most surrogate peptides, the marker/PSA ratios were lower in men with low volume/low grade than those with high volume/high grade; thus, more cancer-associated proteins were produced and secreted into urine from significant cancers.

25   However, most surrogate peptides did not show a significant differentiative power in identifying the low-risk from the high-risk cancers (FIGS. 18A-18B). One exception was MMP9. Two surrogate peptides, FQTFEGDLK and LGLGADVAQVTGALR, produced a *P* value of 0.022 in comparing low volume/low grade cancer and significant cancer (FIGS. 18A-18B and FIG. 19A). The AUC values obtained from an ROC

30   analysis for the two MMP9 peptides were above 0.90; thus, MMP9 and Gleason and

tumor volume were associated in this patient cohort. This result is supported by the

dataset query of cancer cell type transcriptomes. The array signal intensity value for

MMP9 in Gleason pattern 4 cancer cells is 3004.10, 12-fold higher than 238.41 in

Gleason pattern 3 cancer cells. (FIGS. 20A-20B). Other secreted proteins, such as

5   AGR2, showed lower array signal intensity values in Gleason 4 cancer cells than

Gleason 3 cancer cells. For comparison, urine PSA and serum PSA showed no power

in differentiating the two cancer types (FIGS. 19B-19C). This analysis was blinded

because the pathology parameters were not made known before the quantitative SRM

measurements.

10

References

[1] Sokoll LJ, Ellis W, Lange P, et al. A multicenter evaluation of the PCA3 molecular
    urine test: pre-analytical effects, analytical performance, and diagnostic accuracy.
    Clin Chim Acta 2008;389:1-6.

15   [2] van Gils MPMQ, Hessels D, van Hooij O, et al. The time-resolved fluorescence-
    based PCA3 test on urinary sediments after digital rectal examination; a dutch
    multicenter validation of the diagnostic performance. Clin Cancer Res 2007;13:939-
    43.

[3] Hessels D, van Gils MPMQ, van Hooij O, et al. Predictive value of PCA3 in urinary
20   sediments in determining clinico-pathological characteristics of prostate cancer.
    Prostate 2010;70:10-6.

[4] Durand X, Xylinas E, Radulescu C, et al. The value of urinary prostate cancer gene
    3 (PCA3) scores in predicting pathological features at radical prostatectomy. BJU Int
    2012;110:43-9.

25   [5] Day JR, Jost M, Reynolds MA, Groskopf J, Rittenhouse H. PCA3: from basic
    molecular science to the clinical lab. Cancer Lett 2011;301:1-6.

[6] Pascal LE, Goo YA, Vêncio RZN, et al. Gene expression down-regulation in CD90[+]
    prostate tumor-associated stromal cells involves potential organ-specific genes. BMC
    Cancer 2009;9:317.

[7] Pascal LE, Vêncio RZN, Page LS, et al. Gene expression relationship between prostate cancer cells of Gleason 3, 4 and normal epithelial cells as revealed by cell type-specific transcriptomes. BMC Cancer 2009;9:452.

[8] Wayner EA, Quek S, Ahmad R, et al. Development of an ELISA to detect the secreted prostate cancer biomarker AGR2 in voided urine. Prostate 2012;72:1023-34.

[9] Shi T, Gao Y, Quek S, et al. A highly sensitive targeted mass spectrometric assay for quantification of AGR2 protein in human urine and serum. J Proteome Res 2014;13:875-82.

[10] True LD, Zhang H, Ye ML, et al. CD90/THY1 is overexpressed in prostate cancer-associated fibroblasts and could serve as a cancer biomarker. Modern Pathol 2010;23:1346-56.

[11] Bu H, Bormann S, Schäfer G, et al. The anterior gradient 2 (AGR2) gene is overexpressed in prostate cancer and may be useful as a urine sediment marker for prostate cancer detection. Prostate 2011;71:575-87.

[12] Shi T, Su D, Liu T, et al. Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics. Proteomics 2012;12:1074-92.

[13] Shi T, Fillmore TL, Gao Y, et al. Long-gradient separations coupled with selected reaction monitoring for highly sensitive, large scale targeted protein quantification in a single analysis. Anal Chem 2013;85:9196-203.

[14] Shi T, Fillmore TL, Sun X, et al. Antibody-free, targeted mass-spectrometric approach for quantification of proteins at low picogram per milliliter levels in human plasma/serum. Proc Natl Acad Sci USA 2012;109:15395-400.

[15] Shi T, Qian WJ. Antibody-free PRISM-SRM for multiplexed protein quantification: is this the new competition for immunoassays in bioanalysis? Bioanalysis 2013;5:267-9.

[16] Farrah T, Deutsch EW, Omenn GS, et al. State of the human proteome in 2013 as viewed through PeptideAtlas: comparing the kidney, urine, and plasma proteomes for the biology- and disease-driven Human Proteome Project. J Proteome Res 2014;13:60-75.

[17] Nagaraj N, Mann M. Quantitative analysis of the intra- and inter-individual variability of the normal urinary proteome. J Proteome Res 2011;10:637-45.

[18] Fusaro VA, Mani DR, Mesirov JP, Carr SA. Prediction of high-responding peptides for targeted protein assays by mass spectrometry. Nat Biotechnol 2009;27:190-8.

[19] Eyers CE, Lawless C, Wedge DC, et al. CONSeQuence: Prediction of reference peptides for absolute quantitative proteomics using consensus machine learning approaches. Mol Cell Proteomics 2011;10:M110.003384.

[20] MacLean B, Tomazela DM, Shulman N, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics 2010;26:966-8.

[21] Shi T, Sun X, Gao Y, et al. Targeted quantification of low ng/mL level proteins in human serum without immunoaffinity depletion. J Proteome Res 2013;12:3353-61.

[22] Goo YA, Goodlett DR, Pascal LE, et al. Stromal mesenchyme cell genes of the human prostate and bladder. BMC Urol 2005;5:17.

[23] Park K, Tomlins S, Mudaliar K, et al. Antibody-based detection of *ERG* rearrangement-positive prostate cancer. Neoplasia 2010;12:590-8.

[24] Keshishian H, Addona T, Burgess M, Kuhn E, Carr SA. Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. Mol Cell Proteomics 2007;6:2212-29.

[25] Worboys JD, Sinclair J, Yuan Y, Jorgensen C. Systematic evaluation of quantotypic peptides for targeted analysis of the human kinome. Nat Methods 2014;11:1041-4.

[26] Zhang Q, Fillmore TL, Schepmoes AA, et al. Serum proteomics reveals systemic dysregulation of innate immunity in type 1 diabetes. J Exp Med 2013;210:191-203.

[27] Végvári A, Sjödin K, Rezeli M, et al. Identification of a novel proteoform of prostate specific antigen (SNP-L132I) in clinical samples by multiple reaction monitoring. Mol Cell Proteomics 2013;12:2761-73.

[28] Mikolajczyk SD, Marks LS, Partin AW, Rittenhouse HG. Free prostate-specific antigen in serum is becoming more complex. Urology 2002;59:797-802.

[29] Marks LS, Bostwick DG. Prostate cancer specificity of PCA3 gene testing: examples from clinical practice. Rev Urol 2008;10:175-81.

[30] Ho ME, Quek S, True LD, et al. Bladder cancer cells secrete while normal bladder cells express but do not secrete AGR2. Oncotarget 2016; DOI:10.18632/oncotarget.17400.

[31] Nakanishi H, Groskopf J, Fritsche HA, et al. PCA3 molecular urine assay correlates with prostate cancer tumor volume: implication in selecting candidates for active surveillance. J Urol 2008;179:1804-9.

[32] He J, Sun X, Shi T, et al. Antibody-independent targeted quantification of TMPRSS2-ERG fusion protein products in prostate cancer. Mol Oncol 2014;8:1169-80.

[33] He J, Schepmoes AA, Shi T, et al. Analytical platform evaluation for quantification of ERG in prostate cancer using protein and mRNA detection methods. J Transl Med 2015;13:54.

[34] Engholm-Keller K, Larsen MR. Technologies and challenges in large-scale phosphoproteomics. Proteomics 2013;13:910-31.

[35] Udby L, Bjartell A, Malm J, et al. Characterization and localization of cysteine-rich secretory protein 3 (CRISP-3) in the human male reproductive tract. J Androl 2005;26:333-42.

[36] Jiang Z, Fanger GR, Woda BA, et al. Expression of alpha-methylacyl-CoA racemase (P504S) in various malignant neoplasms and normal tissues: a study of 761 cases. Hum Pathol 2003;34:792-6.

[37] Bolduc S, Lacombe L, Naud A, et al. Urinary PSA: a potential useful marker when serum PSA is between 2.5 ng/mL and 10 ng/mL. Can Urol Assoc J 2007;1:377-81.

[38] Nam RK, Saskin R, Lee Y, et al. Increasing hospital admission rates for urological complications after transrectal ultrasound guided prostate biopsy. J Urol 2013;189:S12-7.

[39] Oesterling JE, Jacobsen SJ, Chute CG, et al. Serum prostate-specific antigen in a community-based population of healthy men. Establishment of age-specific reference ranges. JAMA 1993;270:860-4.

In view of the many possible embodiments to which the principles of the disclosure may be applied, it should be recognized that illustrated embodiments are only examples of the disclosure and should not be considered a limitation on the scope of the invention. Rather, the scope of the invention is defined by the following claims. We therefore claim as our invention all that comes within the scope and spirit of these claims.

We claim:

     1.   A method of treating a subject with prostate cancer, comprising:

measuring expression of at least two prostate cancer-related molecules in a sample obtained from a subject, wherein the prostate cancer-related molecules comprise anterior gradient 2 (AGR2), anterior gradient 3 (AGR3), cysteine-rich secretory protein 3 (CRISP3), chemokine (C-C motif) ligand 3 (CCL3), carcinoembryonic antigen-related cell adhesion molecule 5 (CEACAM5), carcinoembryonic antigen-related cell adhesion molecule 6 (CEACAM6), interleukin 24 (IL24), matrix metalloproteinase 9 (MMP9), CXC motif ligand 14 (CXCL14), cluster designation 90 (CD90), periostin (POSTN), and secreted frizzled-related protein 4 (SFRP4);

measuring increased expression of the at least two prostate cancer-related molecules as compared to a control representing expression for each of the at least two prostate cancer-related molecules expected in a sample from a subject who does not have prostate cancer; and

administering at least one of watchful waiting, active surveillance, surgery, radiation, hormone therapy, chemotherapy, brachytherapy, cryotherapy, ultrasound, bisphosphate therapy, biologic therapy, or vaccine therapy to the subject with prostate cancer, thereby treating the subject.

     2.   A method of diagnosing prostate cancer in a subject, comprising:

detecting the expression of at least two prostate cancer-related molecules in a sample obtained from a subject, wherein the prostate cancer-related molecules comprise AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4;

comparing the expression of the at least two prostate cancer-related molecules in the sample obtained from the subject to at least one control representing expression for each of the at least two prostate cancer-related molecules expected in a sample from a subject who does not have prostate cancer; and

determining that the subject has prostate cancer when increased expression of the at least two prostate cancer-related molecules between the sample and the control is detected.

5   3.   The method of claim 1 or claim 2, wherein the at least two prostate cancer-related molecules comprise:

all of AGR2, AGR3, CEACAM5, CD90, and SFRP4; or

all of AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4.

10

4.   The method of any one of claims 1-3, wherein the at least two prostate cancer-related molecules comprise the combinations listed in FIG. 15.

5.   The method of any one of claims 1-4, wherein the at least two prostate

15   cancer-related molecules are low-abundance molecules.

6.   The method of claim 5, wherein the low-abundance molecules comprise AGR2, AGR3, CCL3, CEACAM5, and CEACAM6.

20   7.   The method of any one of claims 1-4, wherein the at least two prostate cancer-related molecules are moderate-to-low-abundance molecules.

8.   The method of claim 7, wherein the moderate-to-low-abundance molecules comprise CRISP3, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4.

25

9.   A method of treating a subject with intermediate- or high-risk prostate cancer, comprising:

measuring expression of MMP9 in a sample obtained from a subject, wherein the expression is determined based on MMP9 protein concentration using at least one

surrogate peptide of MMP9, wherein the surrogate peptide is at least one of
FQTFEGDLK or LGLGADVAQVTGALR;

measuring increased expression of MMP9 in the sample obtained from the
subject as compared to a control representing expression of MMP9 expected in a
5   sample from a subject who has low-risk prostate cancer; and

administering treatment for intermediate- or high-risk prostate cancer, thereby
treating the subject.

10.   A method of diagnosing intermediate- or high-risk prostate cancer,
10   comprising:

detecting expression of MMP9 in a sample obtained from a subject, wherein the
expression is determined based on MMP9 protein concentration using at least one
surrogate peptide of MMP9, wherein the surrogate peptide is at least one of
FQTFEGDLK or LGLGADVAQVTGALR;

15   comparing the expression of MMP9 in the sample obtained from the subject to
expression of MMP9 expected in a sample from a subject who has low-risk prostate
cancer; and

determining that the subject has intermediate- or high-risk prostate cancer when
increased expression of MMP9 between the sample and the control is detected.

20

11.   The method of any one of claims 1-8, wherein the expression is
determined based on protein concentration.

12.   The method of claim 11, wherein the protein concentration is determined
25   using the concentration of at least one surrogate peptide of the protein.

13.   The method of claim 12, wherein the at least one surrogate peptide is a
peptide listed in FIG. 4.

14.   The method of any one of claims 1-13, wherein the protein concentration is determined using an immunoassay.

15.   The method of claim 14, wherein the immunoassay is an ELISA.

5

16.   The method of any one of claims 1-13, wherein the protein concentration is determined using mass spectrometry.

17.   The method of claim 16, wherein the mass spectrometry technique is

10   LC-SRM, LG-SRM, or PRISM-SRM.

18.   The method of any one of claims 1-17, further comprising normalizing expression of the at least two prostate cancer-related molecules to a prostate protein concentration.

15

19.   The method of claim 18, wherein normalizing to prostate protein concentration comprises normalizing to PSA concentration.

20.   The method of claim 19, wherein normalizing to PSA concentration

20   comprises normalizing to the concentration of at least one surrogate peptide of PSA.

21.   The method of claim 20, wherein at least one surrogate peptide of PSA is IVGGWECEK or LSEPAELTDAVK.

25   22.   The method of any one of claims 1-21, wherein the sample obtained from the subject is a urine sample.

23.   The method of any one of claims 1-22, wherein the subject is human.

24.   The method of any one of claims 1-23, wherein the detected expression of the at least two prostate cancer-related molecules has an AUC value of greater than 0.75.

5

PANEL OF PROSTATE CANCER-ASSOCIATED SECRETED PROTEINS FOR PROSTATE DIAGNOSIS AND PROGNOSIS

5   **ABSTRACT OF THE DISCLOSURE**

Methods are provided for treating a subject with prostate cancer and/or diagnosing a subject at risk for prostate cancer, which can include measuring increased expression of at least two prostate cancer-related molecules in a sample obtained from a

10   subject, including the prostate cancer-related molecules AGR2, AGR3, CRISP3, CCL3, CEACAM5, CEACAM6, IL24, MMP9, CXCL14, CD90, POSTN, and SFRP4. The methods can include administering at therapy to a subject with prostate cancer. Methods are provided for treating a subject with intermediate- or high-risk prostate cancer, which can include measuring increased expression of MMP9 in a sample

15   obtained from a subject compared to a control representing expression of MMP9 expected in a sample from a subject who has low-risk prostate cancer.